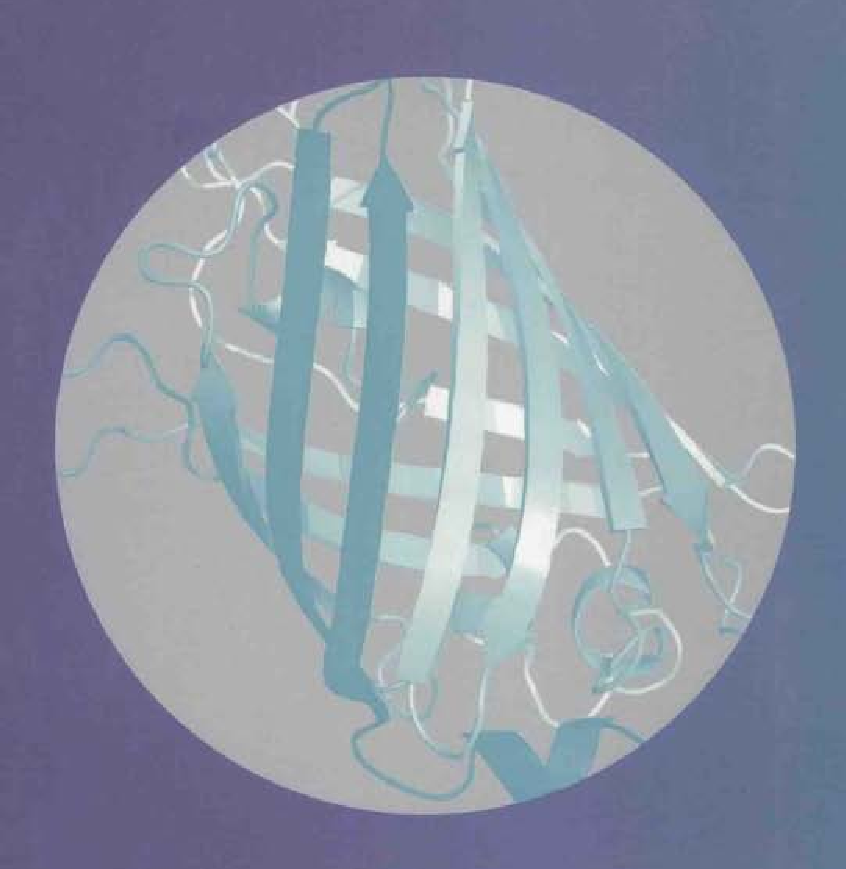


普通高等教育"十一五"国家级规划教材　　普通高等教育精品教材



生物化学原理

Principles of Biochemistry

主编　杨荣武

（第2版）

高等教育出版社

生物化学原理 | 主编 杨荣武

Principles of

Biochemistry



内容简介

数字课程网站

网　址：http://res.hep.com.cn/35696
http://res.hep.edu.cn/35696

数字课程账号：

明码 4130 5241 6614 6261

密码

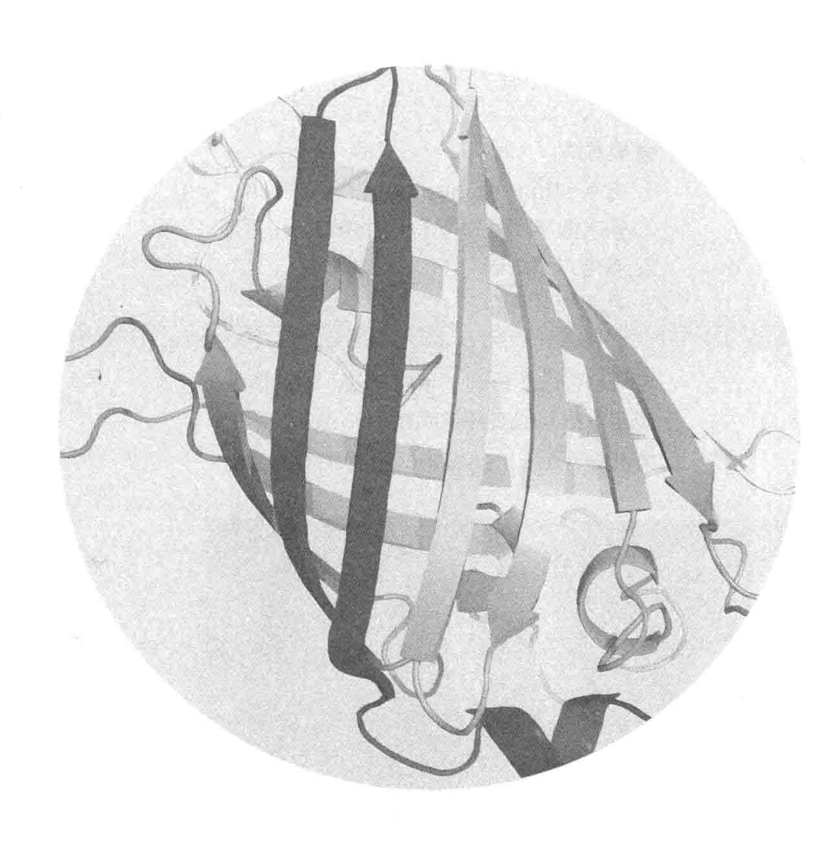
使用说明详见书内数字课程说明页


ISBN 978-7-04-035696-0

定价 88.00 元



普通高等教育"十一五"国家级规划教材　　普通高等教育精品教材



生物化学原理

Shengwu Huaxue Yuanli

（第2版）

主编　杨荣武

参编人员（按姓氏拼音排序）

杜希华　赫福霞

李根喜　戚金亮

沈萍萍　杨荣武

杨永华　郑典元

高等教育出版社·北京
HIGHER EDUCATION PRESS　BEIJING

图书在版编目（CIP）数据

生物化学原理 / 杨荣武主编．—2 版．—北京：高等教育出版社，2012.9

ISBN 978-7-04-035696-0

Ⅰ．①生… Ⅱ．①杨… Ⅲ．①生物化学－高等学校－教材 Ⅳ．①Q5

中国版本图书馆 CIP 数据核字（2012）第 185172 号

策划编辑 王 莉
责任编辑 孟 丽
书籍设计 张申申
责任绘图 尹 莉
责任印制 朱学忠

…………

出版发行 高等教育出版社
购书热线 010-58581118
社　　址 北京市西城区德外大街 4 号
邮政编码 100120
印　　刷 涿州市星河印刷有限公司
开　　本 890 mm × 1240 mm 1/16
印　　张 43
字　　数 1 300 千字
咨询电话 400-810-0598
网　　址 http://www.hep.edu.cn
http://ww

网上订购 http://www.landraco.com
http://www.landraco.com.cn
版　　次 2006 年 10 月第 1 版
2012 年 9 月第 2 版
印　　次 2012 年 9 月第 1 次印刷
定　　价 88.00 元

…………

物料号 35696-00

数字课程（基础版）

生物化学原理

登录方法：

1. 访问 http://res.hep.com.cn/35696
2. 输入数字课程账号（见封底明码）、密码
3. 点击"LOGIN"
4. 进入学习中心

账号自登录之日起一年内有效，过期作废。
使用本账号如有任何问题，
请发邮件至：lifescience@pub.hep.cn

登录以获取更多学习资源！



http://res.hep.edu.cn/35696

第2版前言

经过五年多的努力，《生物化学原理》第2版终将面世，适逢南京大学110周年校庆，就把它作为一份特别的礼物献给我一直深爱的南大吧！

本书第1版有幸入选普通高等教育"十一五"国家级规划教材，并于2007年被评为教育部"普通高等教育精品教材"。第2版在保留先前特色的基础上，做了许多优化、改进和创新，具体包括：

(1) 内容进行了全面的更新。首先更新了每章后面的参考文献和推荐网址，文献引用截至2012年1月，数据引用截至2012年6月。同时根据学科发展更新和添加了一些知识点，例如：在蛋白质结构这一章，添加了兼职蛋白和天然无结构蛋白；在核酸的结构与功能这一章，添加了许多新发现的RNA；在酶学部分，新加入了第十一章"核酶的结构与功能"；在生物氧化这一章，添加了呼吸链复合体Ⅰ产生质子梯度的"分子蒸汽机"的机制；在光合作用这一章，加入了CF_1F_0-ATP合酶的调控机制；在原核生物的基因表达调控这一章，添加了威力强大、可同时对付RNA病毒和DNA病毒等外源核酸序列的CRISPR系统。在分子生物学部分，结合古菌分子生物学机制研究的最新进展，不再按照原核和真核两大系统来组织内容，而是按照细菌、古菌及真核生物三大系统来组织分子生物学"中心法则"的核心内容。

(2) 可读性与易读性进一步提高。为了做到这一点，笔者字斟句酌，反复推敲每个句子，尽可能简明流畅。同时，继续邀请优秀的本科生参与试读教材，充分听取他们的意见，力争使第2版的内容更加生动、深入浅出。例如，有关HbF和HbA在结构和功能上比较的内容，第1版是这样叙述的："胎儿血红蛋白HbF($\alpha_2\gamma_2$)与O_2的亲和力明显高于成人的HbA，这显然有利于它们从母体的胎盘（氧分压大大低于肺）中获取O_2。之所以HbF与O_2亲和力高于HbA，是因为与α亚基结合的γ亚基不能结合BPG。"而第2版却是这样叙述的："胎儿血红蛋白HbF($\alpha_2\gamma_2$)与O_2的亲和力明显高于成人的HbA，这显然有利于胎儿通过胎盘（氧分压大大低于肺）从母亲那里获取O_2。HbF与O_2亲和力之所以高于HbA，是因为与α亚基结合的γ亚基不能结合BPG。由此可见，无私的母爱在分子水平上已经开始了！"。再如，为了提高教学效果，本版改为双色版，突出反应主线，使内容更加鲜活生动。

(3) 调整结构体系，将原来的八篇重组为三篇——结构生物化学、代谢生物化学和分子生物学，由43章扩充到44章。在每篇之前有内容简介和学习方法，还在一些章节内穿插了一些学习方法的介绍，如氨基酸单字母缩写和必需氨基酸的记忆方法，这特别便于读者自学。此外，第三篇分子生物学的内容也做了适当简化。

(4) 引入了许多小框和小测验(Quiz)，小框中有生化动态、身边的生物化学、生化探究、生化趣事、生化传奇、生化与健康和生化小故事等。小框中丰富多彩的内容可进一步激发学生学习生化的兴趣，而小测验可以让读者带着问题去学习。

(5) 特配数字课程(http://res.hep.com.cn/35696或http://res.hep.edu.cn/35696/)，在数字平台上将各章小结设计成填充的形式，以有效地提炼知识点；配有较多拓展内容以及全部小测验的答案，可以协助读者深入学习；提供了4套PPT课件，两套是中文，两套是英文，各分高级版和初级版。相信对于许多学校现在推行的双语教学有所裨益。

(6) 配套的生物化学国家级双语教学示范网站(http://js.nclass.org/vc/802276)和本人的博客(http://blog.sina.com.cn/njuyangsir)已经建成,内有许多辅助教学和学习的资料可供下载。希望有助于读者的学习和教师的教学,力求在读者和编者之间提供了一种相互交流和学习的平台。

(7) 配套的《生物化学学习指南与习题解析》将很快修订更新。

参加第2版编写的有:南京大学生物化学系杨永华教授(参编第四十四章"重组DNA技术")、李根喜教授(参编第四章"蛋白质的性质、分类及研究方法")、沈萍萍教授(参编第十三章"酶的应用及研究方法")、戚金亮副教授(参编第十三章"酶的应用及研究方法"),山东师范大学杜希华教授(参编第十九章"生物能学"),连云港师范高等专科学校郑典元副教授(参编第十八章"代谢总论")以及东北农业大学赫福霞老师(参编第四十四章"重组DNA技术")。本版其他章节则由我编写。

在编写的过程中,对我支持最多的是我的家人。在此,我想要对他们大声地说一声"谢谢"。我要特别感谢我的宝贝女儿——杨诗雅,在本书第1版出版时她还是个中学生,而现在已读大二了。她充分利用业余时间,认真校对了本书。还要感谢南京大学朱煜博士也参与了校对工作,2010级的殷幸煜同学为本书封面提供了良好的设计思路。当然更要真诚感谢我最最可爱的学生们,尤其是2004级邓永昊同学和2005级杨江南同学,他们提供了许多可取的建议。还要特别感谢高等教育出版社对我的信任,以及王莉、孟丽和单冉东等编辑为我提供的所有帮助。同时感谢所有读者对这本书的厚爱。最后,感谢Steve Jobs及时发明了iPad,让我后续的修校工作变得那么方便和轻松。

杨荣武

2012年2月28日于南京

致同行

尊敬的同行,您好!

谢谢您选用本教材。希望这本教材对您的教学会有帮助!同时还特别希望您能对本教材提出宝贵的意见,以便再版的时候进行改进,从而使这本教材越编越好。我的目标是让这本教材早日成为中国最好的中文生物化学教材之一。要想实现这个目标,显然您的帮助和支持是必不可少的。

为了让您能更好地使用本教材,我这里有几点建议,仅供参考。

(1) 课件的选择　与本教材配套的共有4套PPT课件,两套英文版,两套中文版。英文课件用于双语教学。中、英文课件中各有高级版和初级版两种。课时低于70个学时,建议使用初级版;课时在100个学时左右,建议使用高级版。初级版涉及的内容是生物化学中最基础和最重要的内容,高级版则在初级版内容的基础上有很多扩充,在内容的深度和广度都有很大提高。

(2) 课件的使用　在使用PPT课件的时候,我认为最重要的一点是切忌照着PPT一条一条去念,否则会让学生昏昏欲睡,慢慢丧失学习的兴趣! PPT不是万能的,有时候适当地做做板书还是有必要的,甚至可以考虑不定期地进行"裸讲"。

(3) 授课内容的取舍　在使用任何一本教材的时候,上面的内容不可能也不需要都去讲授。在这个学时不断被压缩的年代,对于飞速发展的生物化学来说,在有限的学时内完成高质量的教学是很难的。对此我的建议是:先将教科书上的内容分为课堂上必讲的部分、给学生自学的部分、与其他课程重叠的部分和有兴趣才需要学的部分。教师应该将课本上最难又最重要的部分作为课堂必讲的内容,那些很容易理解的内容可让学生自学,而那些与其他课程重叠的内容事先最好与其他课程的老师进行协调,以避免不必要的重复。像生物化学书上有关氨基酸和单糖的化学反应性质这样的内容,完全可以省去不讲,因为它属于有机化学的内容。酶动力学部分有关多底物反应的动力学可以让有兴趣的学生去自学。在这本教材中我把生物化学分为结构生物化学、代谢生物化学和分子生物学,其中的一个原因就是方便任课老师对上课内容的选择。如果贵校本来就有"分子生物学"这一门课,我建议生物化学课只讲结构生物化学和代谢生物化学就可以了。如果贵校有"植物生理学"这门课,我建议代谢生物化学中的光合作用就可以省去。另外,我还建议,有关实验技术的内容(如层析和电泳)可以直接留给实验课的老师。

(4) 书中小测验(Quiz)的用途　本教材每章都插有富启发性的Quiz,这些Quiz紧扣知识点。我建议可利用这些Quiz在课堂上组织学生进行讨论。为了吸引学生能够主动参与讨论,可将其作为平时成绩的一部分。我的方法是:课堂参与有5分奖励,作为100分以外的附加分,每参与一次课堂讨论就统计一次,最后参与次数最多的就可得5分附加分,其他人根据次数的多少,得相应的附加分。

(5) 小框和科学故事的使用　本教材中有许多小框和科学故事,内容十分有趣,估计学生课后自己会看。但单是看看是不够的,可以让学生选择一个科学故事来写一点心得体会,或者根据最新的科学进展来创作小框故事。这样做可更好地激发学生的学习兴趣,也可以培养学生的科研思维。

(6) 书中文献的使用　本教材每章后面都有参考文献,有经典的,也有最新的。如何使用这些文献呢? 我建议有选择地让学生去阅读,如Watson和Crick有关DNA双螺旋结构的文献是一定要让学

生去阅读的，若是双语教学，可让学生将阅读的文献翻译成中文，并提供机会让学生之间进行文献阅读和翻译的交流。

(7) 第一堂生物化学课怎么上　第一堂生物化学课非常重要，直接关系到学生对这门课的兴趣。因此，老师一定要在第一次课上动用一切可用的手段，激发学生学习生物化学的兴趣。另外，在第一次课上也要把评分标准、课程要求和纪律讲清楚。有的老师喜欢在第一堂课上大谈生物化学的历史，我认为并不适合。与其大谈遥远的生物化学史，不如多讲些最新的进展和突破。这里，我想把我每学期第一次课讲授的主要内容与各位同行分享一下：

- 致欢迎辞，播放一曲"Biochemistry"。
- 让一位同学谈谈自己对生物化学这门课的认识。
- 介绍课堂要求和课堂纪律，公布评分标准。
- 介绍学好生物化学的技巧(可请上一届学得比较好的学生到现场介绍学习心得)。
- 介绍近几年诺贝尔化学奖和诺贝尔生理学或医学奖有关生物化学的内容。
- 介绍当年和以前的由《科学》杂志评选的年度分子或年度突破。
- 介绍什么是生物化学及其主要内容和用途。
- 播放三段视频：第一段是一位伟大的母亲为了治疗孩子的囊性纤维变性，自学生物化学，最后发现一种治疗方法的故事；第二段是一名杜克大学的本科生在被诊断为脊索癌以后，自学生物化学并对这种疾病进行研究的故事；第三段是由哈佛大学制作的模拟细胞内发生的生物化学分子事件。
- 布置课后作业。

(8) 如何进行生物化学课的双语教学　现在提倡双语教学，但做好并不容易。我校的生物化学课现在是国家级双语教学示范课程，对此我的建议是：不主张用全英文授课，因为生物化学的专业词汇太多，使用全英文上课对学生的要求太高。但使用全英文的PPT是完全可行的。至于讲课的英文比例，可控制在50%左右。对于第一次遇到的专业词汇，在讲英文以后，最好立刻快速翻译成中文；课后可适当地补充一些英文阅读资料(如www.sciencedaily.com每天发布的最新生物化学进展)，或者要求学生阅读章后的英文文献，以提高学生的专业英语水平；免费租借给学生英文原版教材；考试不必出全英文试卷，但可以最后出两道共10分的英文题目作为附加题，并要求学生必须用英文回答。

顺致

教安！

南京大学生命科学学院

杨荣武

2012年2月28日

致学生

亲爱的同学,你好!

非常感谢你使用本教材！希望这本教材会让你喜欢上生物化学,并对你学好生物化学有帮助！同时,还特别希望你能对本教材提出宝贵的意见,以便再版的时候进行改进,从而使这本教材日臻完善,越编越好。我一直认为,你们对一本教材的评价是最有发言权的,因此我特别看重你的建议。请给我的邮箱——robertyang@nju.edu.cn 或 thermogenin@gmail.com 或 askmenow@whoever.com 发邮件吧,或者在我的新浪博客——http://blog.sina.cn/njuyangsir 留言吧！我的目标是让这本教材早日成为中国最好的中文生物化学教材之一。要想实现这个目标,你的帮助和支持是必不可少的。

为了让你能更好地使用本教材,让它更好地为你服务,我这里有几点建议,仅供参考。

(1) 让自己喜欢上生物化学　在学生物化学之前,你可能已从学长那里得到一些有关这门课的信息。你所得到的信息中,最多的可能是这门课有多难学。但你不能因此就不敢去学,或者失去学习的兴趣。我始终认为,要想学好一门课,你首先得让自己喜欢上这门课。当然,让你喜欢上生物化学是要有理由的。我这里至少可以给你三个重要的理由:首先,生物化学很有用。从“生物化学”课中,你可以得到很多与医药、营养、保健、疾病的预防和治疗等有关的知识。这些知识可以让你受用一辈子,而且你也可以将这些生物化学知识传播给你的家人和朋友。各种媒体每天都可能出现夸大甚至虚假的医药、保健品的广告,这些广告利用的就是大众缺乏生物化学知识这一点。从某种意义上,学好了生物化学可以让你成为半个医生。这里请允许我提出若干问题,看你知道不知道——爱斯基摩人为什么少得心血管疾病？正常的人需要补脑黄金(DHA)吗？左旋肉碱真的可以减肥吗？为什么路边的蘑菇你不要乱采？为什么过夜的韭菜不能吃？生物化学里的XP代表什么？骆驼为什么几个月可以不喝水？为什么狗急会跳墙、人急会生智？蜘蛛丝和钢筋相比,哪一个强度更高？你有多少个编码蛋白质的基因？生物化学中的“近朱者赤,近墨者黑”是指什么？有天生不会得艾滋病的人吗？如何自酿葡萄酒和自制酸菜？甲醇和氰化钾中毒后如何解毒？阿司匹林是如何消炎的？这些问题的答案,在这本教科书上你都能找到。其次,学好生物化学是你学好生命科学其他课程的基础,比如细胞生物学、遗传学、分子生物学和植物生理学。还有一点,就是考研和工作的需要。如果你将来要考与生命科学有关的研究生,生物化学是必考的课目。

(2) PPT 课件的使用　建议你事先将 PPT 课件打印出来,上课前能预习一下,听课的时候可以在打印稿上直接做笔记。如果你想同时提高自己的专业英语水平,就打印英文 PPT 吧。专业英语不是你想象的那样难,平时多看看就习惯了。此外,在看英文 PPT 的时候,要注意单词的发音,要特别注意重音的位置。为了方便你自己与外国同行进行学术交流,就大胆地将专业词汇读出来吧!

(3) 教科书内容的取舍　要学好生物化学,平时应该多花点时间去阅读教材。但书上的每一个部分不是都要看的。这取决于你想掌握多少生物化学知识或者你的任课老师要求你掌握多少。当然,你首先应该阅读与老师上课讲到的相关内容,其次阅读你感兴趣的部分。此外,在阅读的时候,最好结合本书中提出来的问题(Quiz),或者自己提出问题。对于书中给出的问题,要尽量自己去解决,如果

解决不了，再去问老师，或者去网上查阅答案。

(4) 小框和科学故事的使用　本教材中有许多小框和科学故事，小框中的内容和科学故事十分有趣，强烈建议利用课余时间琢磨琢磨，以享受学习生物化学给你带来的乐趣。但是，单是看看是不够的，最好看过以后多多思考故事给你带来的启示，更希望你能根据生物化学中其他经典的发现和最新的突破写成你的小框故事。如果你愿意，我可以将精彩的小框故事放到下一版《生物化学原理》中，让更多的人分享和体验。

(5) 书中文献的使用　本教材的每一章后面都有参考文献，有经典的，也有新近前沿的。如何使用这些文献呢？建议你有选择地去阅读，一些特别经典的文献是非读不可的，如 Watson 和 Crick 有关 DNA 双螺旋结构的论文，你也可以找几篇最新的文献好好研读和揣摩一番。

(6) 如何利用书中推荐的网址　本书每一章都精选了若干与此章内容相关的网址，而在全书的最后也推荐了 25 个有用的网站。建议你有空就去看一看，这些网站对你的生物化学学习肯定有很大帮助。其中的维基百科（http://en.wikipedia.org）和每日科学新闻（http://www.sciencedaily.com），建议你日常浏览，它们对拓宽你的知识面是非常有帮助的。

(7) 如何使用与这本教材配套的《生物化学学习指南与习题解析》　为了更好地让你能学好生物化学，我还专门编写了一本与本教材配套的《生物化学学习指南与习题解析》。在这本辅助教材中，不仅有很多富于启发性的例题、习题及其解析过程，还有很多有用的学习方法、知识点总结和记忆秘诀。建议你不妨去图书馆借阅或者自己购买一本。

(8) 如何考出好的成绩　每一个学生都想考出好的成绩。对此我的建议是，你首先要搞清楚任课老师的评分标准。请参考评分标准来制定你的学习计划。其次，按照“I hear, I forget; I see, I remember; I do, I understand”，即“听，我忘记了；看，我记住了；做，我理解了”这一句话去做。这句话中的听是听课，看是看书，做是做练习题和做实验。最后，千万不可临时抱佛脚，生物化学内容太多了，平时多看点、多做点吧！

祝生物化学学习顺利、愉快并学业有成！

南京大学生命科学学院

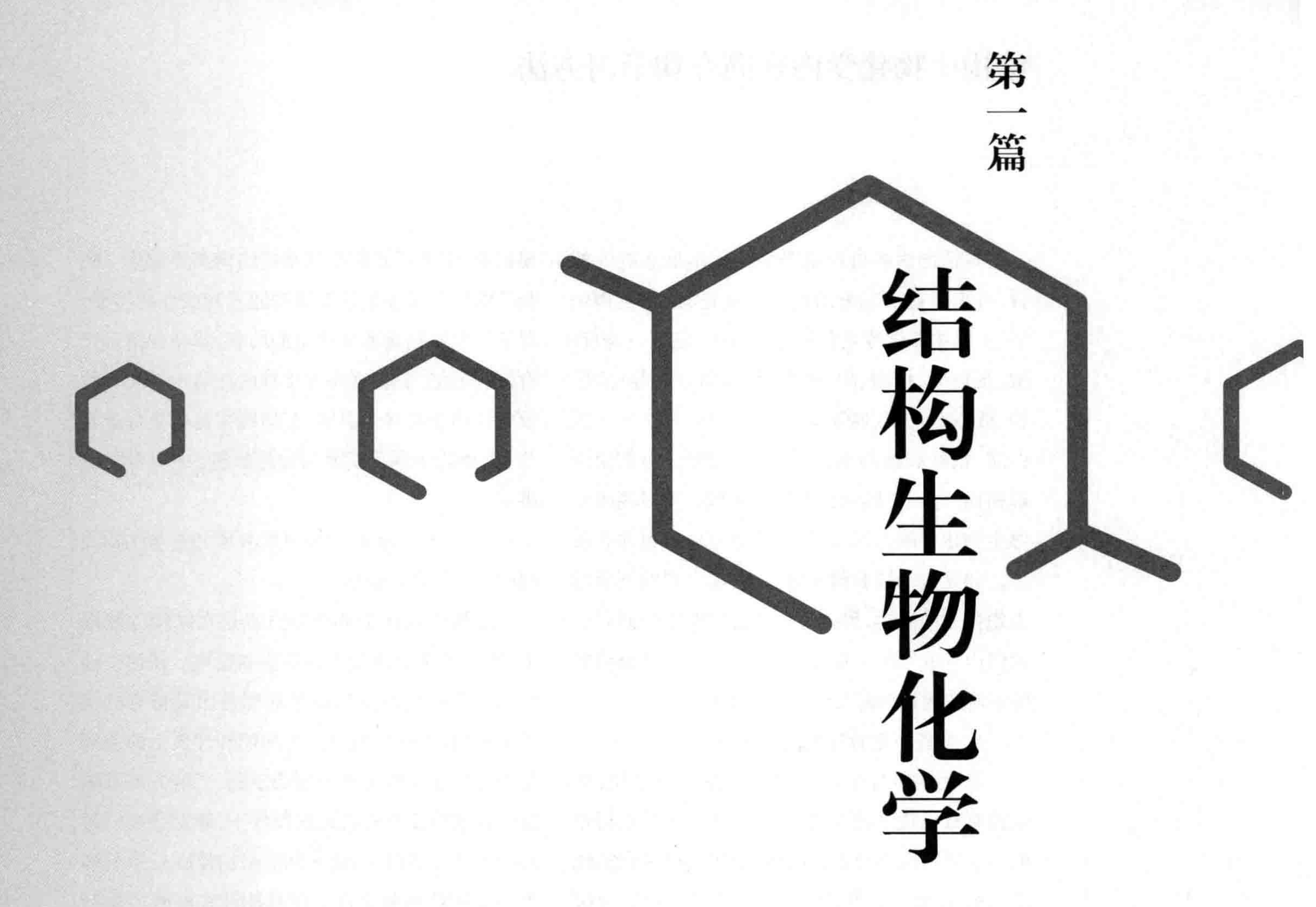
杨荣武

2012 年 2 月 28 日

目录

第一篇

结构生物化学



结构生物化学内容简介和学习方法

本篇的内容可以说是生物化学最基础的部分，有人把这一部分的内容说成是“静态生物化学”。其主要内容是各种生物分子(氨基酸、核苷酸、蛋白质、核酸、酶、糖类、脂质和激素等)的结构、性质与功能，特别是三类生物大分子——蛋白质、核酸和酶的结构、性质与功能。与其他两篇相比，这一篇的内容更容易掌握。但对刚刚学习生物化学的人来说，掌握这部分内容也并非易事。如果这一篇学得不好，一方面会影响到学习生物化学的兴趣，另一方面还会直接影响到以后两篇内容的掌握。那么，如何学好这一部分内容呢？对此我认为需要注意以下几点：

1. 具备一定的有机化学知识

有机化学与生物化学的关系十分密切，如果没有有机化学的基础，想学好生物化学是很难的。在学习生物化学的时候，要用到很多有机化学里的概念和知识，例如手性、构型、构象、双键的顺反、孤对电子、亲电进攻、亲核进攻等。因此，在学生物化学之前，最好先选修有机化学。

2. 对于生物分子的结构，分清哪些需要记，哪些不需要记，而该记的必须想尽一切方法去记住

本篇内容让人感到最头痛的部分是各种生物分子的结构，包括结构式、立体结构、三维结构，以及研究相关结构的方法等，如果是蛋白质和核酸，还有一级结构、二级结构、三级结构，甚至还有四级结构。很多人听到“结构”这两个字就害怕，这是大可不必的。殊不知，并不是每一种生物分子的结构或者每一个层次的结构都需要记。事实上，很多生物分子的详细结构并不需要记，你要用的时候去翻教科书或者有关手册就可以了，而要记的侧重点也会有不同。有的可能仅仅是需要记住一类分子的总体结构，有的则是一种分子结构中最重要的基团。例如，氨基酸的结构通式以及侧链的水溶性(亲水还是疏水)肯定是要记的，但并不需要记住每一种氨基酸侧链基团详细的化学结构。嘌呤环和嘧啶环及其编号需要记的，但并不是所有碱基的结构都要记住。磷脂的结构只需要记住其结构通式和两性的特征。对于一些特别重要又难记的内容，我会介绍行之有效的记忆之法，这些方法有的在与本书配套的数字课程中有详细说明，有的则在与本教材配套的《生物化学学习指南与习题解析》中有专门的讲解。

3. 对于生物分子的功能，要紧扣它与结构之间的关系原则来理解

结构生物化学部分的重点是在结构的基础上，理解各类生物分子的功能和性质。任何一种生物分子所具有的生物学功能是由其特有的结构决定的，不同生物分子在功能和性质上的差别是由它们在结构上的差别决定的。例如，氨基酸的结构性质使其在特定的条件下，能聚合在一起形成多肽和蛋白质，而多肽和蛋白质在三维结构上的多样性和复杂性又使其能够在细胞内执行和完成各种各样的生物学功能；核苷酸的结构性质使其在特定的条件下，能够聚合在一起形成核酸，而核酸的结构特征使其能够充当遗传物质，DNA 与 RNA 在结构上的三大差别使得 DNA 比 RNA 更适合充当遗传物质；单糖的结构特征一方面使其特别适合充当细胞内重要的能源分子，另一方面又使其能够聚合在一起，形成多糖，在结构支持或者能量贮存上起重要作用；磷脂的两性性质使其在水溶液里，能自发组装成生物膜的结构；肌红蛋白和血红蛋白在结构上的共性使它们都能结合氧气，而在结构上的差异使得肌红蛋白适合为肌细胞贮备氧气，血红蛋白则适合为各种细胞运输氧气。因此，在学习生物化学的时候，千万不要孤立地去理解一种生物分子的结构或者功能，而需要将结构和功能有机地结合起来，学会通过观察、分析结构去理解、推测功能，同时也要学会从功能反过来去理解或者还原结构。

4. 注意将各种生物分子的结构、性质和功能进行比较

不同的生物分子在结构、性质和功能上既有

差别，又有某些共同或相似之处。我们在学习的时候，既要掌握不同生物分子之间的差别，又要注意它们在结构和功能上的联系，还要学会分析它们所具有的差别背后的原因和可能的生物学意义。例如，三种生物小分子——氨基酸、核苷酸和单糖，三种生物大分子——蛋白质、核酸和多糖，三种能源贮备——糖原、脂肪和蛋白质，两种核酸——RNA 和 DNA；两类激素——脂溶性激素和水溶性激素，两类生物催化剂——核酶和酶，这几组生物分子我们都可以对它们进行比较。

第一章　氨基酸

氨基酸是一类同时含有氨基和羧基的有机小分子物质。既有氨基又有羧基的特性，使得它们能够彼此缩合成肽，从而作为寡肽、多肽和蛋白质的组成单位。尽管自然界的氨基酸有多种，既有 D 型和 L 型，又有 α 型和 β 型，但是组成蛋白质的氨基酸如果有旋光异构体的话，只能是 L 型的 α–氨基酸。氨基酸除了作为肽的组成单位以外，还是生物体内许多重要活性物质的前体，甚至某些氨基酸本身就具有特殊的生理活性。

本章将主要介绍 22 种蛋白质氨基酸的名称、结构、分类和性质，以及它们重要的生理功能。

第一节　氨基酸的结构、种类和分类

α– 氨基酸的结构通式如图 1–1 所示，其中的 R 表示残余基团（residual group，R）或侧链基团（side chain group），氨基和羧基都与 α– 碳原子相连。不同的氨基酸具有不同的 R 基团，这是区分或分类氨基酸的依据。

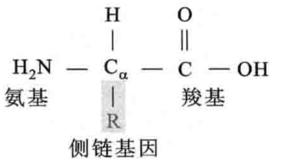


图 1–1　α– 氨基酸的结构通式

一、蛋白质氨基酸

出现在蛋白质分子中的氨基酸称为蛋白质氨基酸（proteinogenic amino acid），又名标准氨基酸（standard amino acid）。蛋白质氨基酸由遗传密码直接决定，目前已发现有 22 种，其中最早发现的 20 种较为常见。

20 种常见的蛋白质氨基酸的名称与结构式见图 1–2至图 1–5。图 1–6 为两种不常见的蛋白质氨基酸，其中含硒半胱氨酸（又称硒代半胱氨酸，selenocysteine）为第 21 种蛋白质氨基酸，只存在于含硒

甘氨酸 (Gly, G)　丙氨酸 (Ala, A)　脯氨酸 (Pro, P)　缬氨酸 (Val, V)　亮氨酸 (Leu, L)　异亮氨酸 (Ile, I)

图 1–2　非极性的脂肪族氨基酸

丝氨酸 (Ser, S)　苏氨酸 (Thr, T)　谷氨酰胺 (Gln, Q)　天冬酰胺 (Asn, N)　甲硫氨酸 (Met, M)　半胱氨酸 (Cys, C)

图 1–3　不带电荷的极性氨基酸

天冬氨酸 (Asp,D)　谷氨酸 (Glu,E)　精氨酸 (Arg,R)　赖氨酸 (Lys,K)　组氨酸 (His,H)

图 1-4　带电荷的极性氨基酸

苯丙氨酸 (Phe,F)　酪氨酸 (Tyr,Y)　色氨酸 (Trp,W)

图 1-5　芳香族氨基酸

含硒半胱氨酸 (Sec,U)　吡咯赖氨酸 (Pyl, O)

图 1-6　第 21、22 种蛋白质氨基酸

蛋白之中，例如谷胱甘肽过氧化物酶（glutathione peroxidase）和一种参与甲状腺素合成的去碘酶。吡咯赖氨酸（pyrrolysine）为第 22 种蛋白质氨基酸，它最为罕见，仅存在于一些能够产生甲烷的原核生物中，作为参与甲烷合成的酶的组分。

22 种蛋白质氨基酸的英文名称可使用三字母或单字母缩写来表示（表 1-1）。有时，在进行氨基酸组成分析的时候，很难区分两种结构紧密相关的氨基酸，例如 Asp/Asn 或 Glu/Gln，这时可以用 Asx 或 B 表示 Asp/Asn，Glx 或 Z 表示 Glu/Gln。

e1-1 氨基酸三字母和单字母缩写记忆秘诀

既然各种氨基酸的差别在于 R 基团，就完全可以根据它的性质来对蛋白质氨基酸进行分类。但对 R 基团的性质可从不同的角度来认定，故分类氨基酸的方法也就不止一种。分类的依据主要有两种：一是根据 R 基团的化学结构，二是根据 R 基团对水的亲和性。

表 1-1 氨基酸的名称、缩写和性质

中文名称	英文全名	三字母缩写	单字母缩写	pK_a(α- 羧基)	pK_a(α- 氨基)	pK_a(R 基团)
丙氨酸	alanine	Ala	A	2.3	9.7	
精氨酸	arginine	Arg	R	2.2	9.0	12.5
天冬酰胺	asparagine	Asn	N	2.2	8.8	
天冬氨酸	aspartic acid	Asp	D	1.9	9.6	3.9
半胱氨酸	cysteine	Cys	C	2.0	10.3	8.4
谷氨酰胺	glutamine	Gln	Q	2.2	9.1	
谷氨酸	glutamic acid	Glu	E	2.2	9.7	4.1
甘氨酸	glycine	Gly	G	2.4	9.8	
组氨酸	histidine	His	H	1.8	9.3	6.0
异亮氨酸	isoleucine	Ile	I	2.4	9.7	
亮氨酸	leucine	Leu	L	2.4	9.6	
赖氨酸	lysine	Lys	K	2.2	9.0	10.5
甲硫氨酸(蛋氨酸)	methionine	Met	M	2.3	9.2	
苯丙氨酸	phenylalanine	Phe	F	1.8	9.1	
脯氨酸	proline	Pro	P	2.0	11.0	
丝氨酸	serine	Ser	S	2.2	9.2	13.6
苏氨酸	threonine	Thr	T	2.1	9.6	13.6
色氨酸	tryptophan	Trp	W	2.4	9.4	
酪氨酸	tyrosine	Tyr	Y	2.2	9.1	10.5
缬氨酸	valine	Val	V	2.3	9.6	
含硒半胱氨酸	selenocysteine	Sec	U	1.9	10.0	5.7
吡咯赖氨酸	pyrrolysine	Pyl	O	未知	未知	

根据 R 基团的化学结构和在 pH 7 时的带电状况,蛋白质氨基酸可以分为 4 类:

(1) 非极性的脂肪族氨基酸　包括 Gly、Ala、Val、Leu、Ile 和 Pro。Gly 是最简单的氨基酸,也是 22 种蛋白质氨基酸中唯一一种没有手性的,其 R 基团仅仅是一个 H 原子;Ala 的 R 基团是非极性的甲基,Val、Leu 和 Ile 的 R 基团则是非极性的带有分支的碳氢链,它们都不能与水分子形成氢键;Pro 实际上是一种亚氨基酸,其侧链与氨基形成一个刚性的环,它和 Gly 一样通常会造成多肽链的弯曲(参看第二章"蛋白质的结构")。

(2) 不带电荷的极性氨基酸　包括 Ser、Thr、Cys、Met、Pyl、Asn 和 Gln。Ser 和 Thr 的 R 基团含有极性的羟基;Asn 和 Gln 分别是 Asp 和 Glu 发生酰胺化反应的产物,它们的 R 基团含有极性的酰胺基。羟基和酰胺基的存在使得这 4 种氨基酸不仅能和水分子,还可以彼此之间或者与肽链骨架、其他极性化合物之间形成氢键。R 基团的亲水性质使得这些氨基酸通常分布在球状蛋白质的表面。Cys 和 Met 的 R 基团都含有 S 原子,其中 Cys 含有巯基,其 pK_a 约为 8.4,因此 Cys 在生理 pH 下主要以非解离的形式存在。溶液中自由的 Cys 分子之间可以通过自发的氧化反应形成二硫键,产生的胱氨酸存在于血液和组织之中,但其水溶性并不高。在蛋白质分子中,两个 Cys 的巯基也可形成二硫键,有助于稳定蛋白质的三维结构。Met 尽管也有 S 原子,但是其侧链基团其余部分是非极性的,因此有人主张将 Met 归为非极性氨基酸。Met 在代谢中的重要作用与 S 原子上的甲基有关,它受到激活以后,可作为甲基供体参与多种生物分子的甲基化修饰,例如 DNA、RNA 和组蛋白的甲基化。

(3) 芳香族氨基酸　包括 Tyr、Phe 和 Trp。它们的 R 基团都含有苯环,但极性差别很大。极性的差别是由苯环上的取代基团造成的。Phe 的苯环上没有取代基团,因此其非极性最强;Tyr 的苯环上含有羟基,能够参与氢键的形成,因此它的 R 基团极性最强;Trp 的环结构较复杂,但其吲哚环上的 N 原子能参与氢键的形成,因而它的极性比 Phe 强。

(4) 带电荷的极性氨基酸　包括 Asp、Glu、Sec、Lys、Arg 和 His，其中前三种带负电荷，后三种带正电荷。Asp 和 Glu 为酸性氨基酸，其 R 基团带有羧基，Sec 的 R 基团是硒醇基（selenol），其 pK_a=5.74，因此在生理 pH 下即发生解离，而带负电荷；Lys、Arg 和 His 为碱性氨基酸，它们的 R 基团在生理 pH 或更低 pH 下被质子化而带正电荷。His 的 R 基团有一个含 N 的咪唑环，Lys 含有的是一级的 ε- 氨基，Arg 含有的是胍基。碱性氨基酸 R 基团上的正电荷能够与带负电荷的基团（如酸性氨基酸的 R 基团和某些辅酶分子上的磷酸基团）形成离子键。此外，酸性氨基酸和碱性氨基酸的 R 基团也参与形成氢键。

如果仅根据 R 基团对水分子的亲和性，可将氨基酸简单地分为两类：

(1) 亲水氨基酸（hydrophilic amino acid）　亲水氨基酸的 R 基团有极性，对水分子具有一定的亲和性，一般能和水分子之间形成氢键。属于亲水氨基酸的有 Ser、Thr、Tyr、Cys、Sec、Asn、Gln、Asp、Glu、Pyl、Arg、Lys、His。

(2) 疏水氨基酸（hydrophobic amino acid）　疏水氨基酸的 R 基团呈非极性，对水分子的亲和性不高或者很低，但对脂溶性物质的亲和性较高。属于疏水氨基酸的有 Gly、Ala、Val、Leu、Ile、Pro、Met、Phe 和 Trp。

在疏水氨基酸之中，Gly 和 Met 的疏水性与其他疏水氨基酸相比要差，原因是 Gly 侧链为氢原子，不是碳氢链，Met 侧链含有与氧原子性质相似的 S，因此有人把这两种氨基酸放到亲水氨基酸之中，或者干脆将它们另立为一类中性氨基酸。

要特别注意的是，所有的氨基酸一般都溶于水，原因是它们都含有亲水的氨基和羧基，只是水溶性有高有低。

R 基团并非是氨基酸分类的唯一标准，有时还可以根据它们对于动物（通常指人）的营养价值，将 20 种常见的蛋白质氨基酸分为必需氨基酸（essential amino acid）和非必需氨基酸（nonessential amino acid）。必需氨基酸是指人体或动物体必不可少，但却不能合成，或者虽能合成，但合成量不够，必须从食物中补充的氨基酸。如果饮食中经常缺少它们，就会影响到机体的健康。必需氨基酸共有 10 种：Lys、Trp、Phe、Met、Thr、Ile、Leu、Val、Arg 和 His。人体虽能够合成 Arg 和 His，但合成的量在特定阶段（如青少年发育和妇女在怀孕期间）不能满足正常的需要，因此这两种氨基酸又称为半必需氨酸（semi-essential amino acid）。余下的氨基酸则属于非必需氨基酸，动物体自身可以进行有效的合成，它们是：Ala、Asn、Asp、Gln、Glu、Pro、Ser、Cys、Tyr 和 Gly。

e1-2 必需与非必需氨基酸名称记忆的秘诀

二、非蛋白质氨基酸

非蛋白质氨基酸也称为非标准氨基酸，在蛋白质生物合成的时候，它们并不能直接参入到肽链之中，要么是蛋白质氨基酸在翻译以后经化学修饰的后加工产物，例如 4-羟脯氨酸（4-hydroxyproline）、5-羟赖氨酸（5-hydroxylysine）和甲酰甲硫氨酸，要么在体内以游离的形式存在，具有特殊的生理功能或者作为代谢的中间物和某些物质的前体，但从来不会参入到蛋白质分子之中，例如在动物体内充当神经递质的 γ-氨基丁酸（gamma-aminobutyric acid，GABA）、作为维生素泛酸组分的 β-丙氨酸和参与尿素循环的鸟氨酸（ornithine）及瓜氨酸（citrulline）（参看第三十一章“氨基酸代谢”）。

第二节　氨基酸的性质和功能

一、氨基酸的性质

氨基酸的性质是由其结构决定的，22 种蛋白质氨基酸都含有 α- 氨基（Pro 为亚氨基）和 α- 羧基，因此有许多共同的性质，而 R 基团的不同则导致个别氨基酸还有某些特殊的性质。

(一) 氨基酸的共同性质

1. 缩合反应

如图 1–7 所示,在一定的条件下,一个氨基酸的氨基可以和另外一个氨基酸的羧基发生缩合反应,以酰胺键或肽键相连形成肽。此反应是肽的人工合成或生物合成的分子基础。

$$H_2N\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}OH + H\underset{H}{\underset{|}{N}}\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}OH \longrightarrow H_2N\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}-\underset{H}{\underset{|}{N}}\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}OH + H_2O$$

肽键

图 1–7 氨基酸的缩合反应——肽键的形成

2. 手性

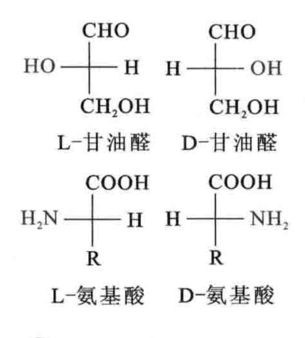


图 1–8 L 型和 D 型氨基酸的 Fischer 投射式结构

除了甘氨酸,其余氨基酸均至少含有一个不对称碳原子(Thr 和 Ile 有两个手性碳原子),因此都具有手性(chirality)。如图 1–8 所示,如果以 L 型甘油醛和 D 型甘油醛为参照物,具有不对称碳原子的氨基酸就有 D 型和 L 型两种对映异构体(enantiomer)或镜像异构体。实验证明,蛋白质分子中的不对称氨基酸都是 L 型。D 型氨基酸仅存在于一些特殊的抗菌肽(如短杆菌肽)和细菌细胞壁的主要成分——肽聚糖之中,它们不能参入到在核糖体上合成的多肽或蛋白质分子之中。

必须指出的是,氨基酸的 D 型和 L 型与氨基酸的旋光方向没有必然的联系,也就是说,某些 L 型氨基酸使偏振光的振动方向顺时针旋转(正值),而另一些则是逆时针旋转(负值)。因此一种氨基酸的旋光方向需要通过实验(旋光仪)来测定。

3. 特殊的酸碱性质与等电点

Quiz1 细菌的肽聚糖分子中含用 D 型氨基酸,这对细菌有什么好处?

氨基酸由于同时含有碱性的氨基和酸性的羧基,因此具有特殊的解离性质,但一种氨基酸的碱性和酸性分别弱于单纯的胺和羧酸。一个氨基酸分子内部的酸碱反应使氨基酸能同时带有正负两种电荷,以这种形式存在的离子称为两性离子(zwitterion)或兼性离子(图 1–9)。实际上,游离的氨基酸在生理 pH 下,主要以两性离子的形式存在。与单纯的胺或单纯的羧酸相比,氨基酸具有更高的熔点(超过 200℃)和更高的水溶性,这是氨基酸主要以两性离子存在的证据。

$$\underset{\text{两性离子}}{H_3\overset{\oplus}{N}\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}O^{\ominus}} + H_3O^+ \underset{}{\overset{\text{在酸性溶液中}}{\rightleftharpoons}} \underset{\text{质子化}}{H_3\overset{\oplus}{N}\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}OH} + H_2O$$

$$\underset{\text{两性离子}}{H_3\overset{\oplus}{N}\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}O^{\ominus}} + OH^- \underset{}{\overset{\text{在碱性溶液中}}{\rightleftharpoons}} \underset{\text{去质子化}}{H_2N\overset{R}{\overset{|}{C}}H\overset{O}{\overset{\|}{C}}O^{\ominus}} + H_2O$$

图 1–9 氨基酸的解离性质以及两性离子的结构

在溶液中,一个氨基酸主要以哪一种形式存在则取决于溶液的 pH。如果不考虑 R 基团可能具有的解离性质,则以两性离子形式存在的氨基酸在强酸溶液中,其羧基接受质子,从而转变为不带电荷的羧基;相反,在强碱性溶液中,其氨基则失去质子而成为不带电荷的基团(图 1–9)。

Quiz2 哪些蛋白质氨基酸的 R 基团可解离? 请用氨基酸单字母缩写来表示。

如果将 R 基团考虑进去,那么有 10 种蛋白质氨基酸的 R 基团具有解离的性质,这些 R 基团是否解离可直接影响到这 10 种氨基酸的带电状态。然而,对于任何一种氨基酸来说,总存在一定的 pH,使其净电荷为零,这时的 pH 称为等电点(isoelectric point, p*I*)。

p*I* 是一种氨基酸的特征常数。当一种氨基酸处于 pH=p*I* 的溶液中,这种氨基酸绝大多数处于两性离子状态,少数可能解离成阳离子和阴离子,但解离成阴、阳离子的趋势和数目相等,由于所带的净电荷为 0,因而若处在电场中,则不会向两极移动。利用上述性质,很容易推导出各种氨基酸 p*I* 的计算公式,也可以使用酸、碱滴定的方法直接测出各种氨基酸的 p*I*(图 1–10)。计算一种氨基酸的 p*I* 的具体步骤是:①找出这种氨基酸所有的可解离基团,并注明它们各自的 pK_a;②假定将它放在极低的 pH 下,这时它所有的可解离基团都处于非解离的质子化状态;③逐步提高溶液的 pH,可解离基团按

框 1–1 生物医药——药物的手性

在 Lewis Carroll 的经典童话书中，女主角爱丽丝一直想知道，镜子里面的牛奶（玻璃奶）是否好喝。对于这个问题，学过生物化学的人一定会这样回答，玻璃奶不一定好喝，但喝下去是一定消化和吸收不了的。

许多有机分子，包括生物分子和药物都有镜像之分，但在生物体内，一般只用或只有其中的一种。例如，氨基酸是 L 型，糖类是 D 型，这是因为催化它们合成以及利用它们的酶具有立体专一性。正因为如此，所以"玻璃奶"中的糖一定是 L 型，而蛋白质中的氨基酸一定是 D 型，这样的牛奶进入消化器官，水解酶识别不了，肠道膜上的转运蛋白也识别不了。

制药工程师们越来越关注制造出来的药品的立体化学纯度，这里举几个例子：1960 年在欧洲，出现了一种处方药物叫沙利度胺，是一种镇静剂。在这一年，数百名在怀孕期间服用沙利度胺的妇女，生下的婴儿有严重的缺陷。原来，她们服用的沙利度胺是两种对映异构体的混合物，其中一种有镇静作用，另一种却可致畸。再如，治疗感冒或过敏的抗组胺药常见的副作用之一是嗜睡，这也是由于抗组胺药的主要成分是对映异构体的混合物，一种导致嗜睡，另一种是很好的减充血剂。目前市场上的止痛药布洛芬也是对映异构体的混合物，但其中的一种更有效。除了药物以外，味觉和嗅觉的产生也具有很强的立体专一性，例如，香芹酮一对对映异构体中的一种具有薄荷气味，另一种具有莳萝的香气；柠檬烯的一对对映异构体中的一种具有柠檬的气味，另一种则具有橙子的香气。药物在体内的生物效应依赖于药物与细胞内特定的酶或蛋白质的相互作用，而味觉和嗅觉的产生则分别依赖于味觉和嗅觉分子与特定受体之间的相互作用，这些相互作用都具有高度的立体专一性。一种药物若是完全用化学手段合成的，最后得到的通常就是消旋体。为了保证药物的安全性，需要将对映异构体分开。

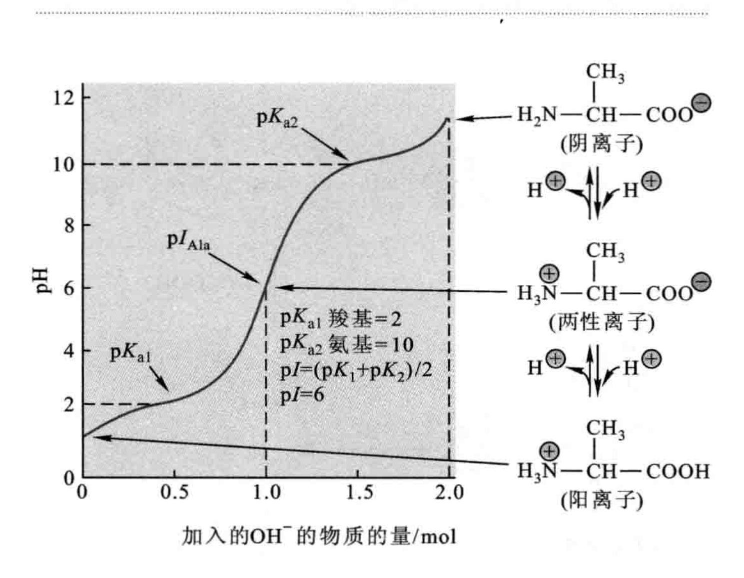


图 1–10 丙氨酸的滴定曲线

照 pK_a 从低到高的顺序依次释放出质子，即 pK_a 越低的就越先释放出质子；④写出所有可能的解离形式，并找出净电荷为 0 的形式；⑤将净电荷为 0 形式两侧的 pK_a 相加除以 2。根据上述计算的方法，不难得出：对于只含有两个 pK_a 的即侧链上无可解离基团的氨基酸（A、F、G、I、L、M、N、P、Q、V、W）而言，它们的 pI 是将两个 pK_a 相加除以 2，对于两个酸性氨基酸（D 和 E）而言，是将两个最低的 pK_a 相加除以 2，而对于 3 个碱性氨基酸（H、K、R）而言，是将两个最高的 pK_a 相加除以 2。

此外，一种氨基酸在 pH 等于其 pI 的溶液中的水溶性最低，因此有些疏水氨基酸在 pH 等于其 pI 的溶液中会发生沉淀。

Quiz3 根据上述方法，计算出 Tyr 和 Cys 的等电点。

4. R 基团的疏水性

R 基团的疏水性（hydropathy）是一种氨基酸的 R 基团对疏水环境的相对亲和能力。一个氨基酸疏水性越高，那么它对疏水环境的亲和力就越高。

氨基酸的疏水性直接影响到蛋白质的折叠。在水溶液之中，疏水氨基酸一般位于蛋白质内部，亲水氨基酸位于蛋白质的表面，这是驱动蛋白质折叠的动力之一。但带相反电荷的极性氨基酸有可能成对地存在于一种球状蛋白质的内部。

5. 氨基酸氨基和羧基参与的化学反应

e1-3 与氨基酸有关的各种反应性质

氨基和羧基都是比较活泼的官能团，它们在特定的条件下，能与多种试剂起反应。这里只介绍几种重要的反应：

(1) 与2,4-二硝基氟苯的反应　在弱碱性溶液中，氨基酸的α-氨基很容易与2,4-二硝基氟苯(2,4-dinitrofluorobenzene，DNFB)起反应，生成稳定的黄色物质——2,4-二硝基苯氨基酸(dinitrophenyl amino acid，DNP-氨基酸)。此反应最初由Frederick Sanger发现，因此也叫Sanger反应，而DNFB也称为Sanger试剂。

多肽或蛋白质在N端游离的α-氨基也能与DNFB反应，但生成的是DNP-多肽或DNP-蛋白质。由于DNP与氨基结合牢固，不易被水解，因此当DNP-多肽被酸完全水解以后，原来的N端氨基酸便成为黄色的DNP-氨基酸。因为DNP-氨基酸溶于乙酸乙酯或乙醚等有机溶剂，所以可以用这两种试剂对其进行抽提，随后进行色谱分析，并以标准的DNP-氨基酸作为对照，就可以鉴定出此氨基酸的种类。Sanger当初就是使用上述方法，测出胰岛素两条链在N端的氨基酸。现在仍然有人使用他发明的方法，来鉴定多肽或蛋白质的N端氨基酸。

$$O_2N-C_6H_3(NO_2)-F + H_2N-CH(R)-COOH \longrightarrow O_2N-C_6H_3(NO_2)-HN-CH(R)-COOH + HF$$

DNFB　　　　DNP-氨基酸

(2) 与异硫氰酸苯酯的反应　在弱碱性条件下，氨基酸的α-氨基可与异硫氰酸苯酯(phenylisothiocyanate，PITC)反应，生成相应的苯氨基硫甲酰氨基酸(PTC-氨基酸)。在酸性条件(如HF和三氟乙酸)下，PTC-氨基酸会迅速环化，形成稳定的苯乙内酰硫脲氨基酸(phenylthiohydantoin amino acid，PTH-氨基酸)(图1-11)。

图1-11　氨基酸与PITC的反应

多肽链N端氨基酸的α-氨基如果没有被封闭，例如甲酰化修饰，那么也能发生此反应，生成PTC-肽。在酸性溶液中，PTC-肽会释放出末端的PTH-氨基酸，而产生比原来少1个氨基酸残基的肽链。新暴露出来的N端氨基可再次进行同样的反应。经过多次重复，N端的氨基酸依次释放出来，成为PTH-氨基酸。由于PTH-氨基酸在酸性条件下极稳定，并可溶于乙酸乙酯，因此在每一次反应结束以后用乙酸乙酯抽提，再经高效液相层析，就可以确定肽链N端氨基酸的种类，并逐步确定出一个完整的多肽链上的氨基酸顺序。氨基酸自动顺序分析仪就是根据该反应原理而设计的。

(3) 与茚三酮的反应　氨基酸与水合茚三酮(ninhydrin)一起在水溶液中加热，可发生反应，生成蓝紫色物质。反应的基本步骤为：首先是氨基酸被水合茚三酮氧化分解成醛、氨和CO_2，同时等量的

水合茚三酮被还原，然后在弱酸性条件下，氨、还原型茚三酮和另一分子茚三酮缩合生成一种蓝紫色物质。

$$H_2N-\underset{\displaystyle R}{CH}-COOH + 2\,\text{茚三酮} \xrightarrow{\text{吡啶}} \text{蓝紫色物质} + CO_2\uparrow + R-CHO + 3H_2O$$

氨基酸　　茚三酮　　蓝紫色物质

所有的氨基酸以及同时具有游离 α- 氨基和 α- 羧基的肽都与茚三酮起反应，并产生蓝紫色物质，只有脯氨酸和它的修饰产物（羟脯氨酸）与茚三酮反应产生黄色物质。此反应十分灵敏，对反应所生成的蓝紫色物质在 570 nm 波长下进行比色法测定，就可确定样品中氨基酸的含量，也可以在分离氨基酸时作为显色剂，对氨基酸进行定性或定量分析。在法医学上，可使用茚三酮反应来采集犯罪嫌疑人在现场留下来的指纹。因为手汗中含有多种氨基酸和肽，遇茚三酮后能起显色反应。

除了以上 3 种重要的化学反应以外，氨基酸还能与亚硝酸、甲醛等发生特征反应，这些性质也可用于氨基酸的定量或定性分析。

e1-4 氨基酸与亚硝酸或甲醛的反应及用途

（二）个别氨基酸的侧链性质和对蛋白质功能的贡献

不同氨基酸的差别在于 R 基团的不同。不同的会给一种氨基酸带来一些特有的性质，同时它们对一种蛋白质的生物学功能也会有不同的贡献。

1. Trp、Phe 和 Tyr 的紫外吸收性质

Trp、Phe 和 Tyr 这三种氨基酸的 R 基团含有苯环，这使得它们对在近紫外（230~300 nm）波长范围内的光具有强吸收（图 1-12）。其中，Trp 和 Tyr 的吸收峰靠近 280 nm，与核苷酸或核酸的吸收不同（吸收峰在 260 nm）。利用上述性质可用来定性、定量测定蛋白质。

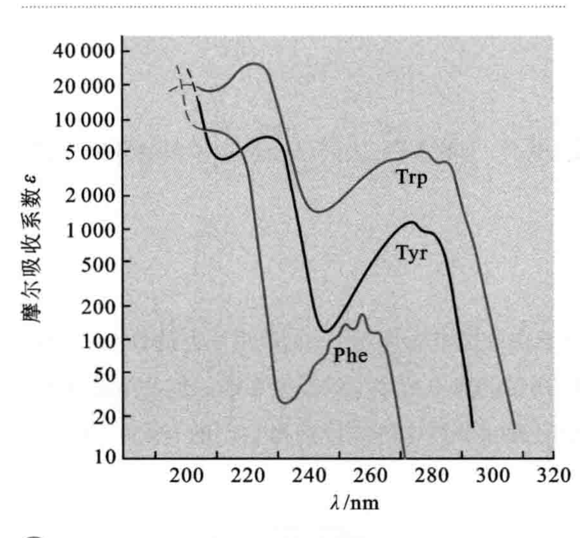


图 1-12　Trp、Phe 和 Tyr 的紫外吸收

2. 个别氨基酸的 R 基团对蛋白质功能的贡献

R 基团有亲水的和疏水的两类，疏水的 R 基团缺乏反应性，因此它们一般对蛋白质的功能没有直接的贡献，所起的作用主要是结构上的。蛋白质需要它们形成疏水核心来驱动折叠，并稳定三维结构，少数蛋白质会利用它们组装成疏水的口袋，以结合脂溶性分子。与此相反，亲水的氨基酸在侧链上含有各种反应性基团，正是这些反应性基团赋予大多数蛋白质能够行使各种各样的生物学功能。例如酶分子，没有亲水基团的参与，它是无法催化反应的。下面就重点解释几种最重要的亲水侧链对蛋白质功能的主要贡献：

（1）含有羟基的侧链　Ser、Thr 和 Tyr 这三种氨基酸的侧链含有羟基，而羟基含有孤对电子，可作为亲核基团参与多种酶的催化。另外，这三种氨基酸侧链上的羟基，可被磷酸化修饰。机体（主要是真核生物）可使用这种方式来调节多种酶和蛋白质的活性。

（2）含有 ε- 氨基的侧链　Lys 的侧链含有 ε- 氨基，而 ε- 氨基含有孤对电子，因此也可以作为亲核基团参与多种酶的催化（详见第十章“酶的催化机理”）。许多蛋白质的辅基含有羧基（如生物素和硫辛酸）能够与 ε- 氨基生成酰胺键，从而实现与蛋白质的共价连接。此外，ε- 氨基可以发生特定的化学修饰，如乙酰化、甲基化和磷酸化。组蛋白可以通过这种修饰调节活性，从而影响到基因的表达（参看第四十三章“真核生物的基因表达调控”）。

(3) 含有巯基和硒醇基的侧链　Cys 含有巯基，巯基也可以作为亲核基团参与多种酶的催化。此外，两个 Cys 的巯基可被氧化成二硫键。二硫键对于许多蛋白质的三维结构有稳定作用(参看第二章“蛋白质的结构”)。Sec 含有更容易氧化的硒醇基，因此它在蛋白质分子中具有抗氧化的活性。

(4) 含有咪唑基的侧链　His 的侧链含有咪唑环，咪唑环的 pK_a 接近 7 的性质使得它在生理条件下，既可以作为质子受体，又可以作为质子供体，故许多酶的活性中心含有 His。

(5) 含有羧基的侧链　谷氨酸和天冬氨酸的侧链含有非 α- 羧基，这种羧基在特定的情况下可以作为质子供体(非解离的状态)或受体(解离的状态)，许多酶利用它们作为一种催化的手段——广义的酸碱催化(参看第十章“酶的催化机理”)。另外，解离的羧基带有负电荷，有些金属蛋白利用它们结合金属离子。

e1-5 与氨基酸 R 基团有关的反应及其用途

此外，某些氨基酸的侧链可以发生特殊的反应，这些反应对于未知氨基酸的鉴定很有用。

二、氨基酸的功能

氨基酸的主要功能包括：①作为各种肽的组成单位；②作为多种生物活性物质的前体，例如，NO 的前体是 Arg，组胺的前体是 His，褪黑激素的前体是 Trp，乙烯的前体是 Met，嘌呤核苷酸从头合成的前体有 Gln、Asp 和 Gly；③作为神经递质，例如，Glu 在脑组织中可作为一种兴奋性神经递质，而它的脱羧基产物 GABA 是一种抑制性神经递质；④其碳骨架可氧化分解产生 ATP；⑤作为糖异生或酮体合成的前体(参看第三十一章“氨基酸代谢”)。

第三节　氨基酸的分离与纯化

根据氨基酸之间在物理性质(大小)或化学性质(电荷)上的差别，人们已经建立了多种分离、纯化氨基酸的方法，其中主要的方法是电泳和层析。

一、电泳

电泳(electrophoresis)是指带电的颗粒或生物分子在外加电场作用下，向带相反电荷的电极做定向移动的现象。显然，带电粒子或分子的大小、形状、所带的净电荷多少等都会影响它们的电泳速率。在待分离样品中，各种生物分子的带电性质以及大小、形状等存在的差异使得它们的“泳动”速率不同，因而可以利用电泳技术对它们进行分离、鉴定或纯化。

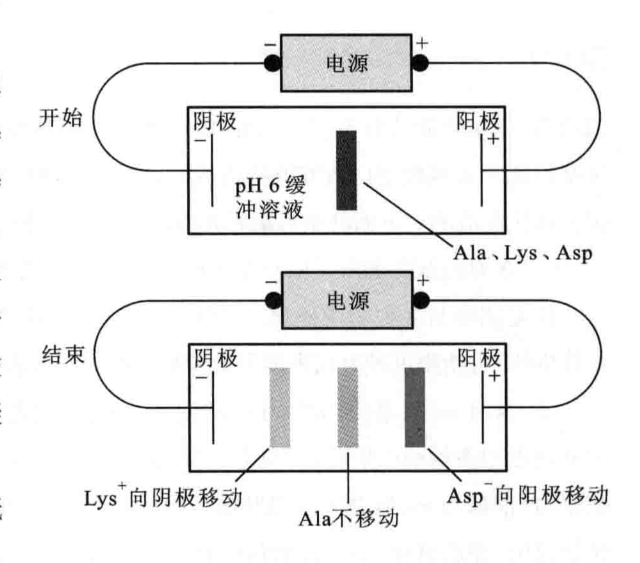


图 1-13　Ala、Lys 和 Asp 的电泳分离

电泳技术有多种方式，但一般根据有无支持物将其分为无支持物的自由电泳(free electrophoresis)和有支持物的区带电泳(zone electrophoresis)两大类。前者包括显微电泳、等电聚焦电泳及密度梯度电泳等。区带电泳则包括以滤纸作为支持物的纸电泳、以醋酸纤维素等薄膜为支持物的薄层电泳以及以凝胶(例如琼脂糖和聚丙烯酰胺凝胶)为支持物的凝胶电泳。

对于小分子的氨基酸来说，比较适合用纸电泳对其进行分离和分析。

以图 1-13 为例，如果将 Ala、Lys 和 Asp 点

在电泳槽中央，在 pH 6 的缓冲溶液中电泳，那么 Lys 和 Asp 因在 pH 6 缓冲溶液中各带正、负电荷会分别向阴极和阳极移动，而 Ala 正好处于 pH 等于其 p*I* 的缓冲溶液中，净电荷为零，所以不发生泳动。

二、层析

早在 1903 年，层析（chromatography）就被俄国植物学家 Mikhail Tswett 用来分离植物色素，因此又称为色谱。所有的层析系统都由两相组成：一个是固定相（stationary phase），它可以是固体物质，也可以是固定在固体物质上的成分；另一个是由可以流动的物质组成的流动相（mobile phase），如水和各种溶剂。当待分离样品随着流动相通过固定相时，各组分在理化性质上的差别使得各自与两相发生相互作用（如吸附、溶解或结合等）的能力不同，最终导致它们在两相中的分配不同，而且随着流动相向前移动，各组分会不断地在两相中进行再分配。与固定相相互作用力越弱的组分，随流动相移动时受到的阻力就越小，向前移动的速率越快；反之，与固定相相互作用越强的组分，向前移动速率越慢。经过分部收集流出液，可得到样品中所含的各单一组分，从而达到分离各组分的目的。

层析的种类繁多，根据流动相和固定相的性质，可以分为液－液层析、液－固层析、气－液层析和气－固层析；根据层析原理分类，又可以分为吸附层析、分配层析、离子交换层析、凝胶过滤层析、亲和层析和疏水层析（参看第四章“蛋白质的性质、分类及研究方法”）；根据操作方式，还可以分为柱层析、薄层层析、纸层析、薄膜层析和高效液相层析（high-performance liquid chromatography，HPLC）。

应用于氨基酸分离的层析主要包括：离子交换层析（ion exchange chromatography）、气相层析（gas chromatography，GC）和 HPLC。这里主要介绍离子交换层析，其他层析的方法将在第四章中做详细的介绍。

离子交换层析是以含有带电基团的树脂做固定相，利用它与流动相中带相反电荷的离子结合并进行可逆交换的性质来分离目标带电分子的一种方法(图 1–14)。如果进行可逆交换的离子是阳离子，则为阳离子交换层析；反之，则称为阴离子交换层析。在进行离子交换层析时，首先需要用几倍于柱体积的低盐缓冲溶液对树脂进行平衡。然后在低盐条件下，将样品上柱。与树脂带相反电荷的分子会通过静电作用与树脂结合，而不带电荷的或带相同电荷的分子直接流出。最后，用高盐缓冲溶液洗脱，以取代与树脂结合的分子。

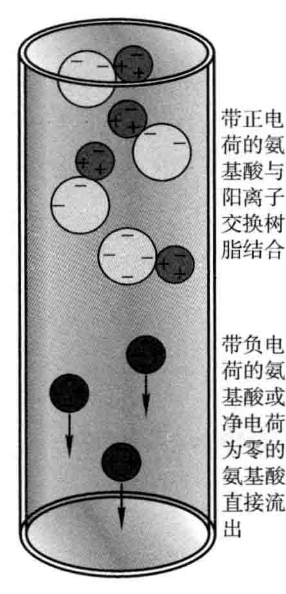


图 1–14　离子交换层析分离氨基酸的原理

同样以 Asp、Lys 和 Ala 三种氨基酸为例。如果选用阳离子交换层析，在 pH 6 的低盐缓冲溶液下上柱，那么带正电荷的 Lys 将被吸附在树脂上，而带负电荷的 Asp 和净电荷为零的 Ala 将直接出现在流出液中，Lys 会在随后的高盐缓冲溶液洗脱中得到，这样就实现了 Lys 与 Asp 和 Ala 的分离。如果还需要将 Asp 和 Ala 分开，可以将收集到的直接流出液再进行阴离子交换层析。

Quiz4 若是使用阴离子交换层析，这三种氨基酸如何分离?

科学故事——第 22 种蛋白质氨基酸的发现

对于生物学家而言，他们因发现新的氨基酸带来的喜悦不亚于物理学家发现了新的粒子，或者化学家发现了新的元素。

1986 年以前，人们一直认为，出现在蛋白质分子中由遗传密码编码的蛋白质氨基酸残基只有 20 种。到了 1986 年，科学家们终于在含硒蛋白中发现第 21 种蛋白质氨基酸——含硒半胱氨酸。时隔 16 年之后，来自美国俄亥俄州立大学的两个研究小组在产甲烷古菌里，发现了第 22 种蛋白质氨基酸——吡咯赖氨酸（Pyl）。

俄亥俄州立大学由 Joseph A. Krzycki 领导的研究小组一直在研究一种产甲烷古菌——巴氏甲烷八叠球菌（*Methanosarcina barkeri*）。此微生物能够将单甲胺（monomethylamine）、二甲胺（dimethylamine）和三甲

胺(trimethylamine)转变成甲烷。1995 年,Krzycki 的研究小组分离到一些与甲烷生成有关的特殊蛋白质。两年以后,他们又分离到编码其中一个蛋白质的基因,并测定出了它的核苷酸序列。1998 年,他们发表了这个基因的全序列,结果显示其可读框内含有一个反常的琥珀型终止密码子(amber codon)。

密码子是决定氨基酸的三字母核苷酸序列(参看第三十九章“蛋白质的生物合成及其在细胞内的降解”),琥珀型终止密码子的核苷酸序列是 TGA,它通常不决定任何氨基酸,它的出现一般是多肽链合成结束的标志。然而,让 Krzycki 吃惊的是该终止密码子竟然编码一种氨基酸,而且这种奇怪的现象还出现在其他几种与产甲烷有关的基因上。与此同时,由 Michael Chan 领导的研究小组开始研究由这个琥珀型密码子编码的氨基酸的结构。他们意识到,这个古怪的密码子也许编码一种新的氨基酸,但也不能排除其他的可能性。Krzycki 及其同事决定测定原来蛋白质的氨基酸序列,当得到蛋白质的氨基酸序列以后,发现由琥珀型密码子决定的氨基酸似乎仅仅是一个赖氨酸。但是,Krzycki 仍然要求 Chan 和他的一个博士研究生 Bing Hao 对含有这个氨基酸的蛋白质晶体结构进行研究,以确定那个氨基酸的性质。经过两年的研究,Chan 和 Hao 终于确定了这个蛋白质的结构。他们得到的一些数据表明那个氨基酸是一个新的氨基酸。

不仅如此,Krzycki 还在寻找其他证据。他和他的几个博士研究生很快发现了一种新的特异性 tRNA,专门负责将新的氨基酸插入到蛋白质之中,同时还发现其他一些与此过程有关系的酶。综合两个发现以及获得的更为详细的晶体结构,他们确信发现了第 22 种由 DNA 编码的蛋白质氨基酸——吡咯赖氨酸。

他们还认为,这个新的氨基酸是一个非常罕见的氨基酸,因此这么多年后才被发现。然而,Krzycki 相信它也可能存在于其他生物体内。Chan 对此表示同意并指出,发现第 22 种氨基酸将激励更多的研究者去寻找第 23 种乃至第 24 种蛋白质氨基酸。随着更多种生物的基因组序列被破译,有理由认为某些有趣的事情迟早会发生。

根据 Krzycki 在 2011 年 3 月 31 日在 *Nature* 上发表的题为“The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine”的论文,现在已很明确,Pyl 是由两个 Lys 转变而来的。

e1-6 本章小结填填看

推荐网址:

1. http://en.wikipedia.org/wiki/Amino_acid(英文维基百科有关氨基酸的内容,既全又新)
2. http://themedicalbiochemistrypage.org/amino-acids.html(完全免费的医学生物化学课程网站,内容丰富、全面,更新快)
3. http://www.biology.arizona.edu/biochemistry/problem_sets/aa/aa.html(美国亚利桑那大学生物化学课程网站,其一大亮点是可进行自测)
4. http://www.bio.fju.edu.tw/excel/content05/html/30.htm(台湾省辅仁大学生物化学课程网站,其在氨基酸命名上有别于大陆)

参考文献:

1. Krzycki J A, et al. The complete biosynthesis of the genetically encoded amino acid pyrrolysine from lysine. *Nature*, 471: 647–650, 2011.
2. Rother M and Krzycki J A. Selenocysteine, pyrrolysine, and the unique energy metabolism of *Methanogenic Archaea*. *Archaea*, 2010.
3. Arnér E S. Selenoproteins-What unique properties can arise with selenocysteine in place of cysteine? *Exp Cell Res*, 316(8): 1 296–1 303, 2010.
4. Hao B, et al. A new UAG-encoded residue in the structure of a methanogen methyltransferase. *Science*, 296: 1 462–1 466, 2002.

第二章　蛋白质的结构

蛋白质是由多个氨基酸通过肽键连接而成的多聚物。其结构具有一定的层次，一般包括一级结构、二级结构、三级结构和四级结构，其中二级结构、三级结构和四级结构统称为高级结构，一级结构决定高级结构。此外，细胞内的蛋白质并不是孤立存在的，而是一起构成蛋白质组。蛋白质组也不是一成不变的，而是处在动态的变化之中。

本章从肽和肽键谈起，将重点介绍蛋白质的各级结构特征、多肽链的折叠机制以及测定和预测蛋白质结构的方法，同时还会对蛋白质组以及蛋白质组学的概念做简单的介绍。

第一节　肽的结构

肽（peptide）是氨基酸之间通过 α-氨基和 α-羧基缩合以酰胺键（amido bond 或 amido linkage）或肽键（peptide bond）相连的聚合物，它包括寡肽（oligopeptide）、多肽（polypeptide）和蛋白质。构成肽的每一个氨基酸单位称为氨基酸残基(residue)。各氨基酸残基以肽键相连形成的链状结构称为肽链(peptide chain)。

一、肽的分类和命名

肽的分类和命名的主要依据是氨基酸残基的数目、组成和排列顺序。

（一）肽的分类

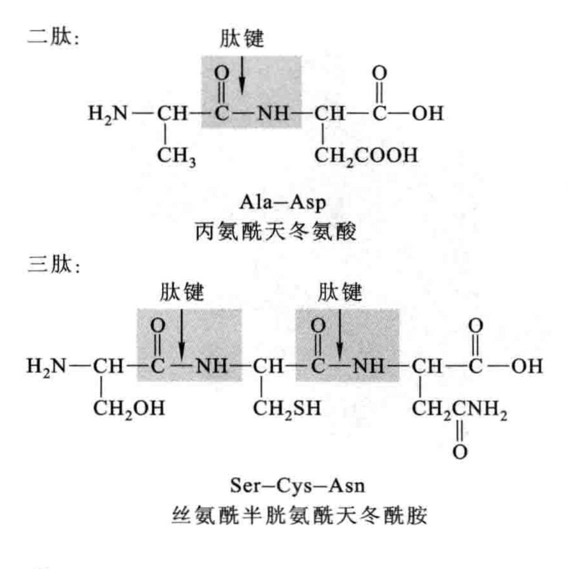


图 2-1　二肽和三肽的结构（N 端→C 端）

肽的分类可根据氨基酸残基的数目而直呼其为几肽。例如，2 个氨基酸形成的肽称为二肽，3 个氨基酸构成的肽称为三肽，依此类推（图 2-1）。一般将由 2~10 个氨基酸残基组成的肽称为寡肽，由 11~50 个氨基酸残基组成的肽称为多肽，由 50 个以上的氨基酸残基组成的肽称为蛋白质。

除了少数环状肽链以外（如 α- 鹅膏蕈碱），其他肽链都含有不对称的两端：其中含有游离的 α- 氨基的一端称为氨基端（amino terminal）或 N 端，含有游离的 α- 羧基的一端称为羧基端（carboxyl terminal）或 C 端。但有一些小肽为了提高稳定性，N 端的氨基和 / 或 C 端的羧基会发生特定形式的化学修饰而被封闭。例如，N 端发生甲酰化或焦谷氨酰化（N 端谷氨酸残基的 γ- 羧基与其 α- 氨基形成酰胺键），C 端发生酰胺化。

（二）肽的写法

按照惯例，书写一条肽链的序列时总是从 N 端到 C 端，即 N 端放在最左边，C 端放在最右边。有时为了强调，可在 N 端和 C 端分别添加 H 和 OH，以此来表示两端游离的氨基和羧基。各氨基酸残基可用三字母或单字母缩写表示。图 2-1 中的三肽可写为 Ser—Cys—Asn、H—Ser—Cys—Asn—OH 或

SCN。

（三）肽的命名

寡肽通常根据其氨基酸残基从 N 端到 C 端的依次排列，被称为某氨基酰某氨基酰……某氨基酸。按照上述规则，图 2-1 中的二肽和三肽应分别命名为丙氨酰天冬氨酸和丝氨酰半胱氨酰天冬酰胺。

多肽和蛋白质的命名一般是根据它们的来源或 / 和功能进行的，如胰高血糖素是一种来源于胰岛，功能是使血糖升高的多肽；血红蛋白是来源于红细胞与运输氧气有关的蛋白质。

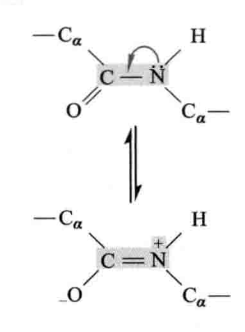


图 2-2　肽键的双键性质

二、肽键的结构与性质

无论是寡肽、多肽还是蛋白质，都是由氨基酸残基通过肽键连接而成，因此首先需要对肽键的结构和性质有所了解。

肽键实际上是一种酰胺键，一般用羰基 C 和酰胺 N 之间的单键表示。肽链中的每一个酰胺基称为肽基（peptide group）或肽单位（peptide unit）。研究表明肽键具有以下性质：

（1）具有部分双键的性质(40%)　其键长为 0.133 nm，短于一个典型的单键，长于一个典型的双键。肽键所具有的双键性质是酰胺 N 上的孤对电子与相邻羰基之间发生共振作用造成的（图 2-2）。

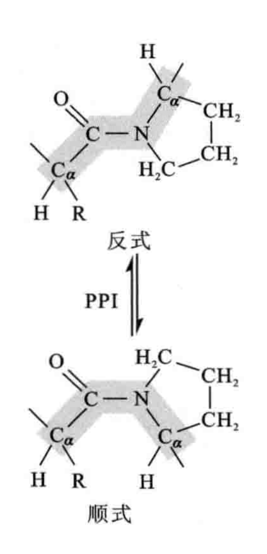


图 2-3　X—Pro 之间的反式肽键和顺式肽键

（2）多为反式（*trans*）构型，也有顺式（*cis*）构型　在反式构型中，邻近的非成键原子的空间位阻更小，因而比顺式构型稳定（相差 100 倍）。如果肽键是由一种氨基酸的羧基与 Pro 的亚氨基形成的（X—Pro），那么仍然是反式的构型稳定，但是 Pro 残基的四氢吡咯环造成的空间位阻会部分抵消反式构型原有在空间位阻上的优势，这使得顺式与反式在稳定性上的差距有所降低（仅相差 4 倍），因此这时候的肽键也有可能以顺式形式存在（图 2-3）。然而，多肽或者蛋白质在核糖体上合成的时候，最初形成的肽键都是反式的，只是后来某些部位的肽键在肽基脯氨酰顺反异构酶（peptidyl prolyl *cis-trans* isomerase，PPI）的催化下变成了顺式。

（3）形成肽平面　与肽键相关的 6 个原子共处于一个平面，此平面结构称为肽平面（peptide plane）或酰胺平面（amide plane）（图 2-4）。

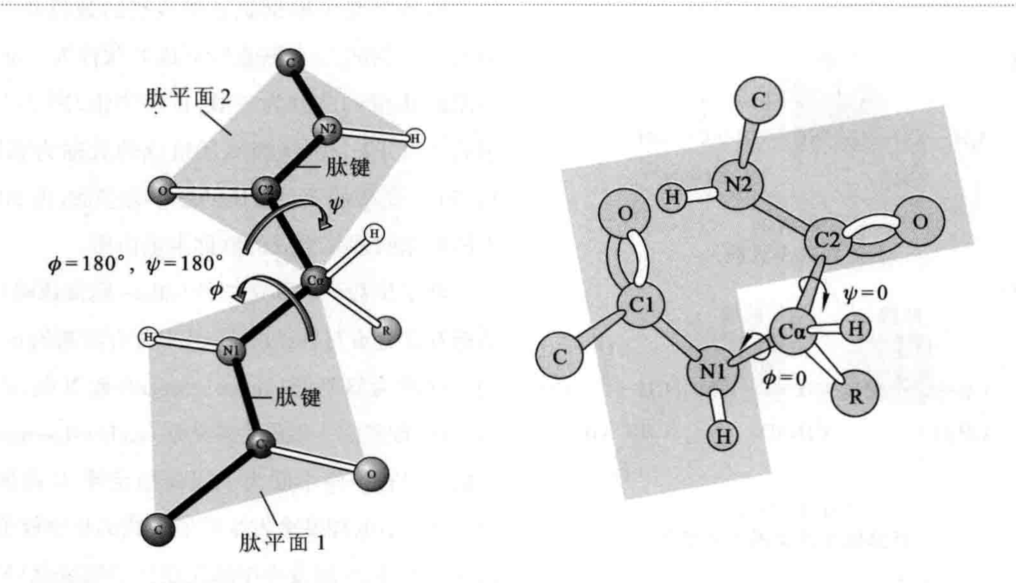


图 2-4　蛋白质分子上的肽平面和二面角（ϕ，ψ）（Campbell 等，2009）

肽平面结构的形成是由于肽键具有双键的性质。在一个肽平面上，有两个与 Cα 有关的单键，这两个单键是可以自由旋转的。正因为如此，每一个肽平面有两个可以旋转的角度：由 Cα—N 单键旋转的角度称为 ϕ（phi），Cα—C 单键旋转的角度称为 ψ（psi）（图 2-4），与同一个 Cα 有关的一对 ϕ 和 ψ 称为蛋白质的二面角（dihedral angle）。当一条肽链上所有的二面角都被确定以后，该肽链的三维结构也就基本确定了。

根据惯例，如果肽链处于完全伸展状态，即所有的肽键位于同一个平面上，ϕ 和 ψ 定为 ±180°；如果 ϕ 的旋转单键 Cα—N1 两侧的 N1—C1 和 Cα—C2 呈顺式时，则规定 ϕ=0°；如果 ψ 的旋转键 Cα—C2 两侧的 Cα—N1 和 C2—N2 呈顺式时，则规定 ψ=0°；从 Cα 向 N1 观察，顺时针旋转 Cα—N1 键得到的 ϕ 角为正值，反之为负值；从 Cα 向 C2 观察，顺时针旋转 Cα—C2 键得到的 ψ 角为正值，反之为负值。

理论上 ϕ 和 ψ 可以是 -180°~180° 的任意值，但由于主链上的原子和侧链基团之间存在空间位阻(steric hindrance)，因而有些角度是不允许的，例如 ϕ 和 ψ 不能为 0°。据估计，蛋白质分子中能出现的二面角仅占理论值的 10% 左右。不同氨基酸残基的侧链基团不一样，由此产生的空间位阻也不一样。Gly 的侧链基团最小，由它产生的空间位阻也最小，与它相关的二面角变化的范围最大。而 Pro 的刚性吡咯环则使其受到的限制最为严重，其 ϕ 值被限制在 -85°~-35°。

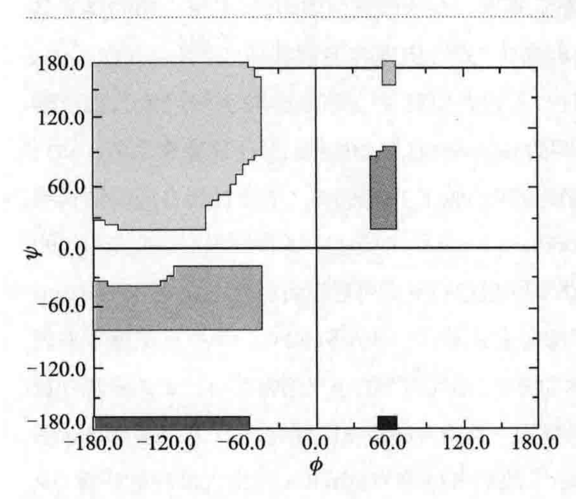


图 2-5 根据计算得到的 Ramachandran 图

为了能够更好地显示一种蛋白质分子所能采取的二面角的值，印度生物化学家 G. N. Ramachandran 通过计算蛋白质中非键合原子间的最小接触距离，预测了哪些成对二面角(ϕ,ψ)所规定的两个相邻肽平面的构象是允许的，哪些是不允许的，并提出以 ϕ 为横坐标、ψ 为纵坐标，在坐标图上标出每一对(ϕ,ψ)作为图中的一个点，以此来表示一种蛋白质的构象。这种作图方法后来被称为 Ramachandran 作图法。

图 2-5 是 Ramachandran 预测到的结果，其中有色区域表示的是所有允许的二面角，白色区域表示的是不允许的二面角。上述结果与后来实验测定出来的结果基本一致。

(4) 酰胺 N 带部分正电荷，羰基 O 带部分负电荷　这也是酰胺 N 上的孤对电子与相邻羰基之间发生共振作用而造成的(见图 2-2)。

三、寡肽的理化性质

寡肽常见的理化性质包括旋光性、两性解离、双缩脲反应和水解反应，其中旋光性和两性解离是由组成的氨基酸直接带来的。

(1) 旋光性　一种寡肽只要不是全部由 Gly 组成，它就具有旋光性，其旋光度应该是各个氨基酸旋光度的总和。

(2) 两性解离　除了含有可解离的侧链基团以外，多数肽还含有自由的 N 端氨基和 C 端羧基，这些可解离的酸性和碱性基团同样赋予肽两性离子的性质。特别是一些寡肽的很多性质与自由的氨基酸相近，例如它们的滴定曲线就很相似。但随着可解离基团的增加，滴定曲线变得越来越复杂。显然肽也有等电点，其中小肽 pI 的计算方法与氨基酸相似，但复杂的寡肽与多肽、蛋白质一样都不能计算得出，只能使用实验的方法(如等电聚焦)进行测定。

Quiz1 为什么多肽和蛋白质的 pI 不能通过计算的方法得到?

(3) 双缩脲反应　双缩脲在碱性溶液中能与硫酸铜反应生成紫红色络合物，此反应称为双缩脲反应(biruet reaction)。凡大于二肽的肽因为含有与双缩脲结构相似的肽键，所以都能发生此反应。但二肽不行，因为它只有一个肽键。

(4) 水解反应　寡肽里面含有肽键，肽键在特定的条件下会发生水解。

四、几种天然存在的活性肽

天然存在的寡肽通常具有特殊的生理活性，因此一般被统称为活性肽。活性肽又可分为两类：一

框 2-1 生化点滴——五花八门的蛋白质命名方法

迄今为止,科学家们已发现几万种不同的蛋白质,而且现在每天都可能有新的蛋白质被发现。对于蛋白质的命名,科学界并没有一种统一的系统命名法则,主要是由发现者来决定,而发现者在命名的时候通常各行其是,随意性很强。那么,蛋白质的命名究竟有没有什么规律呢?

如果一种蛋白质是酶,它的命名就很容易,只要按照酶命名的基本规则进行就可以了,如乙醇脱氢酶、胃蛋白酶和蛋白质激酶 A。

如果发现的蛋白质不是酶,则命名的方法就主要有以下几种:①根据蛋白质的来源、性质、功能或者将它们相结合起来命名,如胰岛素、胶原蛋白、血红蛋白、肌红蛋白、孔蛋白(porin)、绿色荧光蛋白(green fluorescent protein,GFP)、枯草杆菌素、蓖麻毒素和霍乱毒素。在这些蛋白质中,很多蛋白质的英文名称中有后缀“in”,如胰岛素为 insulin。“in”最初被翻译为“素”,但现在多被翻译为“蛋白”,如 cyclin 起初被翻译成周期素,但现在已统一被称为周期蛋白,不过有少数蛋白质的名字仍然带“素”,如泛素(ubiquitin)。②根据蛋白质的大小来命名,一般用小写的 p(protein)加上其相对分子质量来命名。p53、p21 和 p19 是三种重要的抗癌蛋白,可抑制细胞周期的前进而防止细胞癌变。③根据蛋白质的结构特征来命名,如网格蛋白(clathrin)、叉头盒 O 蛋白 1(forkhead box O1,Foxo1)。④根据蛋白质突变引起的疾病来命名。例如,多种参与哺乳动物核苷酸切除修复的蛋白质突变可导致着色性干皮病(xeroderma pigmentosum,XP),因此这些蛋白质用 XP 再加字母编号来表示,像 XPA 和 XPC。⑤用英文缩写来表示蛋白质的名称,这是现在表示蛋白质的一种最简单的方法,相关缩写的字母通常都有一定的来源和意思。例如,参与信号转导的 Sos 蛋白最初是在研究果蝇的一条信号通路中发现的,在一种叫 *sevenless* 基因的下游起作用,发现者把它比作是“son of sevenless”,故后来被称为 Sos;Ras 是指大鼠肿瘤病毒(rat sarcoma virus)编码的一种蛋白质,Rho 是后来被发现的 Ras 的同源物(Ras homologue,Rho);Arf 是一种 ADP 核糖基化因子(ADP-ribosylation factor,Arf);Sar 是分泌相关联的 Ras 家族相关蛋白(secretion-associated and Ras superfamily-related protein);Ran 是指 Ras 相关的 RNA 运输蛋白(Ras-related RNA transport,Ran);Rap 是 Ras 相关蛋白(Ras-related protein,Rap),参与高尔基和内质网之间的膜泡运输。⑥根据发现者或神话中的人物来命名蛋白质。例如,Kornberg 酶是指 Arther Kornberg 发现的 DNA 聚合酶 I,参与真核细胞双链断裂修复的主要蛋白质 Artemis 是希腊神话中一个神的名字。

这些五花八门的蛋白质命名给学习带来了一定的困难,但我们能做的就是尊重发现者的权利;同时这也告诉我们,如果你想将来有机会用自己的方法来命名一种新的蛋白质,那你就得快点发现它。

Quiz2 假定还原型的谷胱甘肽各解离基团的 pK_a 与游离的氨基酸相同,试计算它的 pI。

e2-1 常见的几种活性肽的来源与功能

类是在核糖体上合成,由 DNA 上特定的基因编码。但在核糖体上合成出来的通常是一个多肽前体,需要通过切割才能得到具有特定生物活性的寡肽产物,如催产素就属于此类。在核糖体上合成的肽中不可能参入 D 型氨基酸,且形成的肽键总是由 α- 氨基和 α- 羧基缩合而成;另一类活性肽不是在核糖体上合成的,如谷胱甘肽(glutathione)和短杆菌肽 S。这一类肽在基因组 DNA 分子上并没有自己的基因,可能含有 D 型氨基酸;含有的肽键不一定是由 α- 氨基和 α- 羧基缩合而成的,例如谷胱甘肽的第一个肽键。

第二节 蛋白质的结构

蛋白质的结构一般具有 4 个层次,即一级结构、二级结构、三级结构和四级结构(图 2-6),但并不是所有的蛋白质都具有三级结构或四级结构。

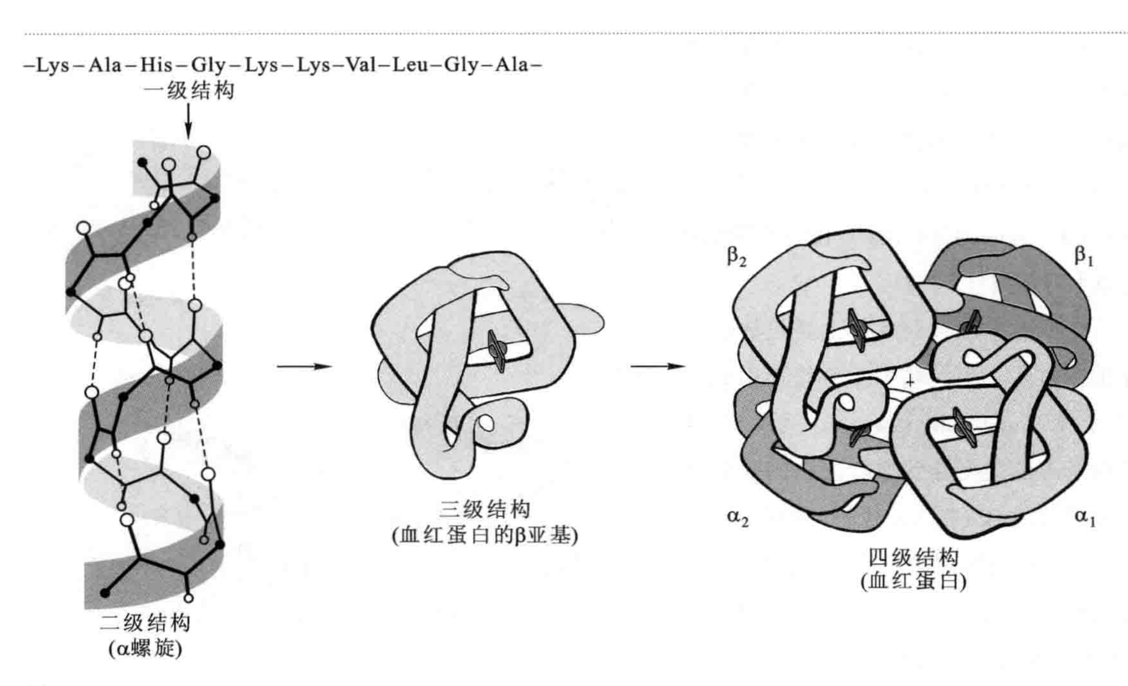


图 2-6 蛋白质的 4 种结构层次(Voet 等,2011)

一、蛋白质的一级结构

蛋白质的一级结构(primary structure)也叫蛋白质的共价结构,是指氨基酸在多肽链上的排列顺序。如果一种蛋白质含有二硫键,那么其一级结构还包括二硫键的数目和位置。

1954 年,Sanger 成功地测定出牛胰岛素(insulin)的一级结构(图 2-7)。胰岛素成为世界上第一例一级结构被测定的蛋白质,Sanger 也因此荣获 1958 年的诺贝尔化学奖。

1965 年,我国科学家根据胰岛素的一级结构,人工合成了具有正常生理活性的牛胰岛素,这是世界上第一种人工合成的蛋白质,从某种意义上来讲,合成的成功也直接证明了 Sanger 测定出的胰岛素的一级结构是正确的。

测定蛋白质一级结构的方法有直接测定法和间接测定法,详见第四章第三节"蛋白质一级结构的测定"。

确定蛋白质的一级结构有助于理解其高级结构和功能,因为一种蛋白质的一级结构包含了决定其高级结构的所有信息,而高级结构又与蛋白质的功能直接相关。来源于不同种类生物的同源蛋白具有相似的氨基酸序列(如细胞色素 c),而具有相似序列的不同肽链往往意味着它们具有相似的二级

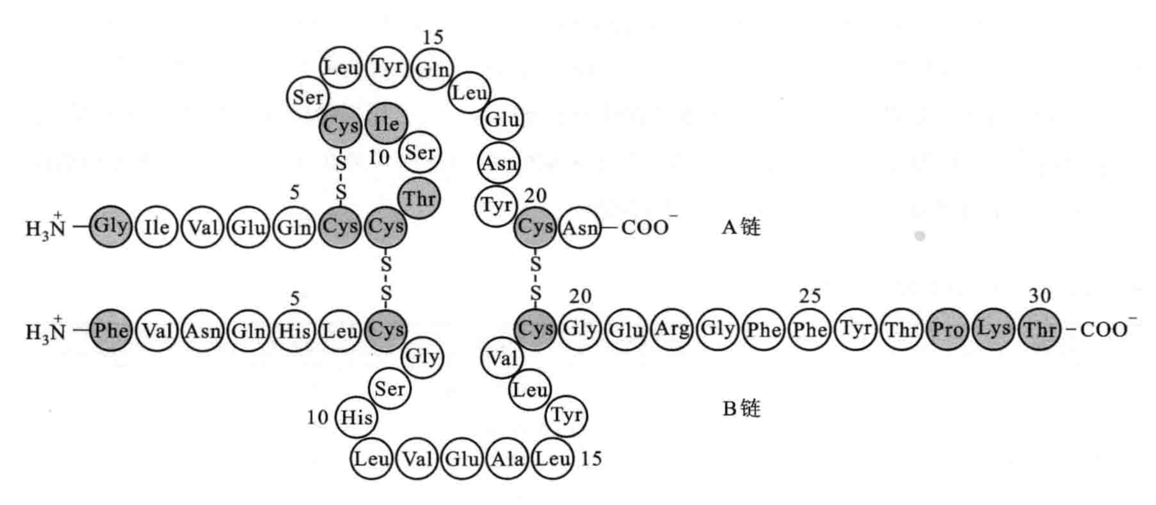


图 2-7 牛胰岛素的一级结构

结构和三级结构。例如，构成血红蛋白的 α- 珠蛋白和 β- 珠蛋白有 46% 的序列是相同的，它们的高级结构也非常相似。再如，人 α- 乳清蛋白（lactalbumin）和鸡溶菌酶有 39% 的序列是相同的，它们的高级结构也十分相似（图 2-8）。

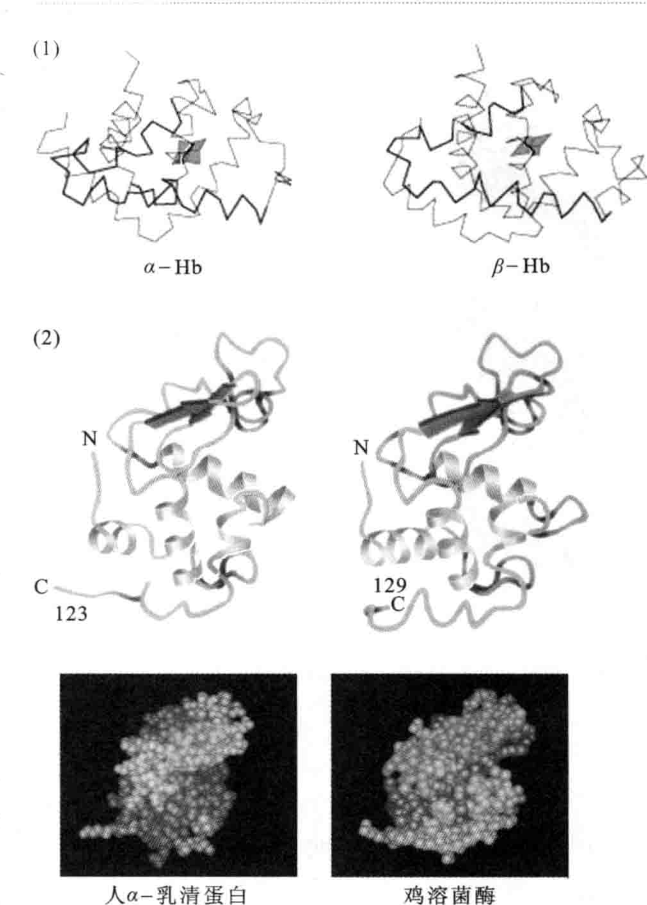


图 2-8 α- 珠蛋白、β- 珠蛋白、人 α- 乳清蛋白和鸡溶菌酶的三维结构

许多遗传性疾病实际上是机体内某种蛋白质的一级结构发生差错而引起的“分子病”（molecular disease）。例如，β- 珠蛋白基因的突变导致血红蛋白的 β 亚基在 6 号位的 Glu 被 Val 代替，带有这种突变的人表现为镰状细胞贫血，这是一种最常见的血红蛋白分子病（参看第三章“蛋白质的功能及其与结构之间的关系”）。

测定蛋白质的一级结构，还可以帮助科学家在分子水平上研究生物的进化。例如，将不同生物来源的细胞色素 c 的一级结构进行比对，可有助于了解物种间的进化关系，因为物种间的亲缘关系越接近，细胞色素 c 的一级结构就越相似（参看第三章“蛋白质的功能及其与结构之间的关系”）。现在因特网上有许多在线的蛋白质序列数据库，如 ExPASy（http://au.expasy.org）、Owl（http://www.bioinf.man.ac.uk/dbbrowser/OWL）和 PCM（http://searchlauncer.bcm.tmc.edu/multi-align/multi-align.html）。感兴趣的人可以利用这些数据库，对已发布的各种蛋白质的一级结构进行比对。

二、蛋白质的二级结构

蛋白质的二级结构（secondary structure）是指多肽链的主链（backbone）部分（不包括 R 基团）在局部形成的一种有规律的折叠和盘绕，其稳定性主要由主链上的氢键决定。

常见的二级结构有 α 螺旋（alpha-helix）、3_{10} 螺旋、三股螺旋（triple helix）、β 折叠（beta-sheet）、β 转角（beta-turn）、β 凸起（beta-bulge）、环和无规则卷曲（random coil）。其中前 6 种比较有规律，所涉及的肽段上的每一对二面角都差不多（表 2-1），在 Ramachandran 图上，则表现为所有的二面角集中在某一区域（图 2-9），而环和无规则卷曲的二面角则落在其他的允许区。从结构的稳定性上看，右手 α 螺旋 >β 折叠 >β 转角 > 无规则卷曲和环；但从功能上看正好相反，酶与蛋白质的活性中心通常由无规则卷曲和环充当，α 右手螺旋和 β 折叠一般只起支持作用。

Quiz3 你如何将环与 β 转角区分开来？试计算它的 pI。

表 2-1 几种二级结构的二面角

结构	ϕ	ψ	结构	ϕ	ψ
完全伸展	−180°	180°	左手 α 螺旋	60°	−60°
反平行 β 折叠	−139°	135°	3_{10} 螺旋	−49°	−26°
平行 β 折叠	−119°	113°	π 螺旋	−57°	−40°
右手 α 螺旋	−57°	−47°	三股螺旋	−51°	153°

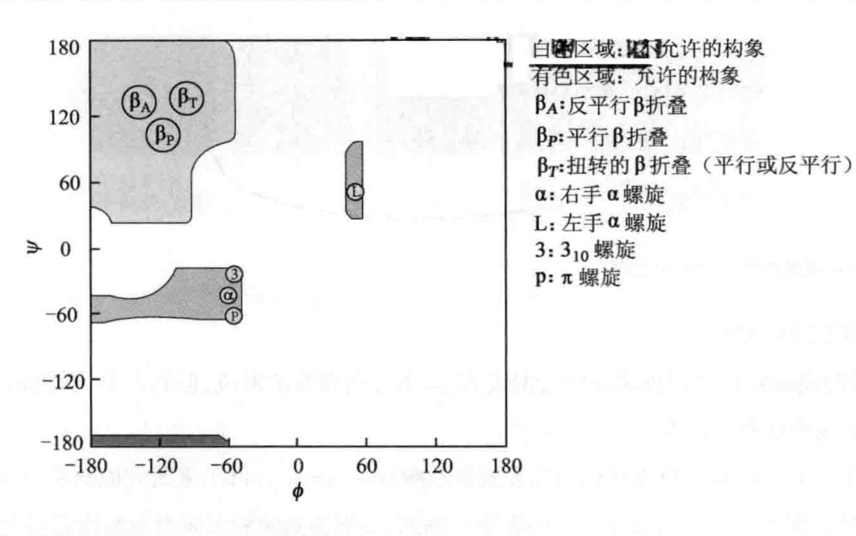


图 2-9 各种二级结构在 Ramachandran 图中的分布

（一）α 螺旋与 3_{10} 螺旋

α 螺旋是一种最常见的二级结构（图 2-10），最先由 Linus Pauling 和 Robert Corey 于 1951 年提出。其主要内容包括：

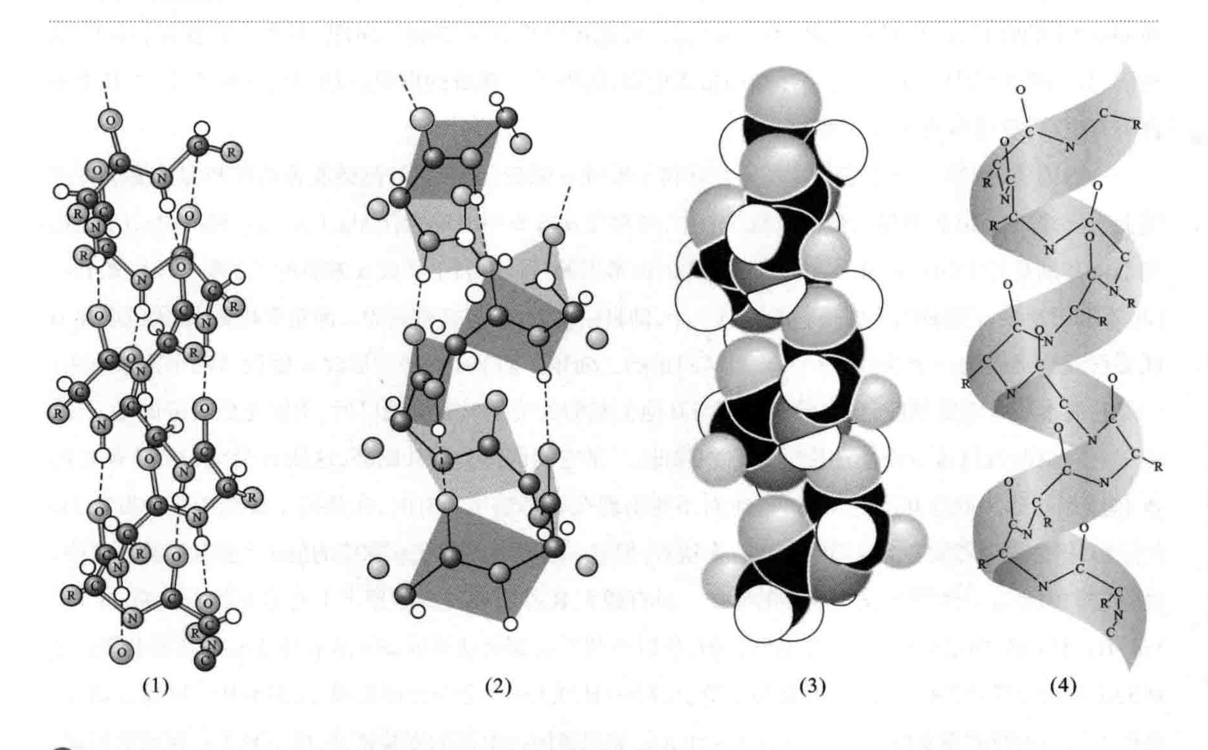


图 2-10 4 种不同方式显示出的 α 螺旋结构以及判断螺旋方向的方法

（1）肽链主链围绕一个虚拟的轴以螺旋的方式伸展。

（2）螺旋的形成是自发的，主要由主链上的氢键来稳定。氢键在 n 位氨基酸残基上的 C=O 与 n+4 位氨基酸残基上的 N—H 之间形成（图 2-11）。被氢键封闭的环共含有 13 个原子，因此 α 螺旋也称为 3.6_{13} 螺旋。螺旋的前、后 4 个氨基酸残基通常不能形成全套螺旋内氢键，这些残基需要与水分子或蛋白质内部的其他基团形成氢键后才能稳定下来。

（3）每隔 3.6 个氨基酸残基，螺旋上升一圈。每一个氨基酸残基环绕螺旋轴 100°，螺距为 0.54 nm，即每个氨基酸残基沿轴上升 0.15 nm。螺旋的半径为 0.23 nm，二面角（ϕ，ψ）约为（−57°，−47°）。

（4）α 螺旋有左手和右手之分，但在蛋白质分子中发现的 α 螺旋主要是右手螺旋，左手螺旋很少见。这是因为蛋白质中的氨基酸只有 L 型，若形成左手螺旋，L 型氨基酸的 β 碳和羰基氧在空间上会

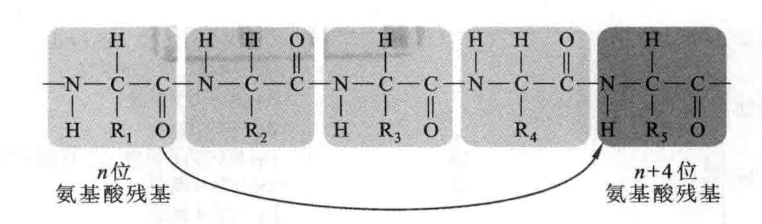


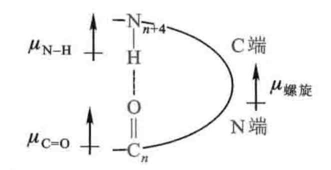
图 2-11　维系 α 螺旋的氢键供体和受体

Quiz4　如果构成一段 α 螺旋的氨基酸是 D 型，那么该段螺旋的方向如何？

发生冲突，其稳定性会降低。

(5) 氨基酸残基的 R 基团伸展在螺旋的表面，虽不参与螺旋的形成，但其大小、形状和带电状态却能影响到螺旋的形成和稳定性。

可以使用由同一种氨基酸残基构成的多聚氨基酸(polyamino acids)，来研究侧链基团的大小、形状和带电状态对 α 螺旋的形成及其稳定性的影响。例如，多聚亮氨酸和多聚丙氨酸很容易形成 α 螺旋；多聚甘氨酸只会形成无规则卷曲；多聚脯氨酸不形成 α 螺旋，但会形成它所特有的右手螺旋(ϕ 和 ψ 分别是 −75° 和 160°，肽键为顺式)或左手螺旋(ϕ 和 ψ 分别是 −75° 和 150°，肽键都为反式)。多聚天冬氨酸和多聚谷氨酸在 pH≥7 的时候，只会形成无规则卷曲；在 pH< 2.5 的时候，则自发地形成 α 螺旋。这是因为如果 pH≥7，Asp 和 Glu 的 R 基团均带负电荷，静电的排斥导致无法形成稳定的 α 螺旋；但在 pH<2.5 的时候，R 基团质子化，不再带电荷，因此可以形成 α 螺旋。同理，多聚赖氨酸在 pH<11 的时候，只形成无规则卷曲，是因为 R 基团带正电荷，阻碍了 α 螺旋的形成；而如果 pH≥12，其 R 基团不再带有电荷，就很容易形成 α 螺旋。



一般而言，判断一个氨基酸残基是否有利于形成 α 螺旋，要看它的侧链能否遮盖和保护螺旋在主链上的氢键。在 20 种常见的蛋白质氨基酸中，经常在 α 螺旋中出现的有 Ala、Cys、Leu、Met、Glu、Gln、His 和 Lys，特别是其中的 Lys、Met、Ala、Leu 和 Glu；最常见的两种不利于形成 α 螺旋的氨基酸是 Gly 和 Pro。Gly 不利于形成 α 螺旋的原因是它的侧链太小，即自由度太大，与它有关的二面角变化太大，无法满足 α 螺旋形成所必需的一个条件——具有相对固定的二面角。而 Pro 不利于形成 α 螺旋与两个因素有关：一是它的本质是亚氨基酸，当它的亚氨基与其他氨基酸的羧基形成肽键以后，不能充当氢键供体，于是主链上氢键的数目减少而使螺旋的稳定性降低；二是它的侧链为刚性的环，这使 α 螺旋形成所需要的 ϕ 值难以达到。但这并不意味着 Pro 绝对不能出现在 α 螺旋中。有时，在某些 α 螺旋中也会发现 Pro 的踪影，例如视紫红质 7 次跨膜的 α 螺旋中就有，但它一般只会出现在 α 螺旋的前 4 个氨基酸残基之中，并且它的出现会导致螺旋发生小小的弯曲。具有较大 R 基团或者 β 碳原子上有分支的氨基酸，如 Ile、Val、Thr、Phe 和 Trp，因为容易产生空间位阻，所以不利于 α 螺旋的形成；Ser、Asp 和 Asn 在侧链上靠近主链的位置含有氢键供体或受体，容易与主链上的 N—H 或 C═O 竞争形成氢键，故也不利于 α 螺旋；肽链上带有电荷的氨基酸，如 Glu、Asp、Lys 和 Arg，如果带同种电荷的连续排列，也不利于 α 螺旋的形成。

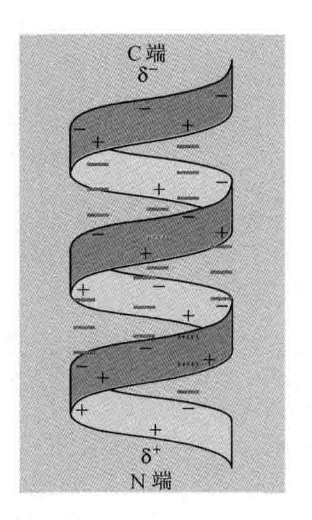


图 2-12　螺旋的总偶极矩

除了 R 基团的大小、形状和带电状态之外，螺旋具有的偶极矩也能影响到螺旋的稳定性。由于稳定 α 螺旋的氢键几乎都平行于螺旋轴，使得每一个肽键的固有偶极"连为一体"，螺旋的总偶极矩(dipole moment)成为一个净偶极矩，其正极在 N 端，负极在 C 端(图 2-12)。虽然偶极矩的存在不利于螺旋的稳定，但蛋白质可以借助于螺旋本身 N 端的酸性氨基酸(R 基团带负电荷)和 C 端的碱性氨基酸(R 基团带正电荷)来中和偶极矩在螺旋两端带有的相反电荷，从而消除其对螺旋稳定性的不利影响。此外，还可以通过相邻螺旋之间的相互作用(两个螺旋反平行排列，相互中和各自在两端的电荷)来降低偶极矩对螺旋的去稳定作用。有些蛋白质通过 α 螺旋在 N 端带有的正电荷与带负电荷的配体分子(如磷酸基团)结合，或者通过在 C 端带有的负电荷与带正电荷的配体(如金属离子)结合，来消除偶极矩对螺旋的稳定性产生的不利影响。

Quiz5　以下 4 段 α 螺旋在相同的条件下，哪一个最稳定，哪一个最不稳定？为什么？
(1) KAAAAAAK
(2) EAAAAAAE
(3) KAAAAAAE
(4) EAAAAAAK

对于一个稳定的 α 螺旋而言，主链上的原子埋在螺旋的内部，彼此之间通过范德华力结合，里面

几乎已没有任何空隙，而所有的 R 基团都呈辐射状离开螺旋的核心，充分展示在螺旋的表面。螺旋的这种性质，方便了它们与其他部位的氨基酸残基的侧链或者与其他生物分子（如 DNA）之间的相互作用。

螺旋的上述性质可以用螺旋轮（helical wheel）作图表示。一个螺旋轮是螺旋沿着螺旋轴的二维投射图，具体做图步骤是：首先想象有一条与螺旋轴平行、垂直于纸平面的直线（0°）。其顶端（正对读者）表示螺旋的第一个氨基酸的侧链，用画圈的数字 1 表示它的位置；然后，按顺时针方向转 100°画第二条直线，使其一端在螺旋轴，另一端用画圈的数字 2 表示第二个氨基酸；再从第二条直线向顺时针旋转 100°画第三条直线，依此类推，直到将螺旋上的所有氨基酸都画上。注意，离我们越近的数字圆圈轮要画得越大。按照这种作图方法，第五圈以后，图的样式将会重复，即第 19 号位的氨基酸残基将位于 1 号位残基的上面。

使用螺旋轮作图能显示出各氨基酸残基侧链相对于螺旋边缘的分布情况，有利于判断一个 α 螺旋是亲水的，还是疏水的，或是两亲的（图 2-13）。如果螺旋上的亲水氨基酸较多，则螺旋可视为亲水的；如果全部是疏水氨基酸，螺旋则是疏水的。疏水 α 螺旋经常出现在膜内在蛋白的跨膜区。例如，在真核细胞的质膜上，有一类叫 G 蛋白偶联的受体（参看第十七章“激素及其受体介导的信号转导”），这类受体属于跨膜蛋白，共跨膜 7 次。每一个跨膜的肽段形成的都是疏水 α 螺旋；如果亲水残基和疏水残基含量差不多，并且亲水的分布在螺旋一侧，形成亲水的一面，疏水的集中在另一侧，组成疏水的一面，那么这样的螺旋称为两亲螺旋（amphipathic helix）。两亲螺旋的疏水面通常与蛋白质内部的疏水区结合，而亲水面与水相互作用。此外，在许多载脂蛋白分子中还发现两亲螺旋，载脂蛋白通过两亲螺旋的疏水面与脂质结合，亲水面则暴露在水溶液之中。

Quiz6 哪些氨基酸残基最适合在跨膜的 α 螺旋的两端，即在水相和油相的交界处？

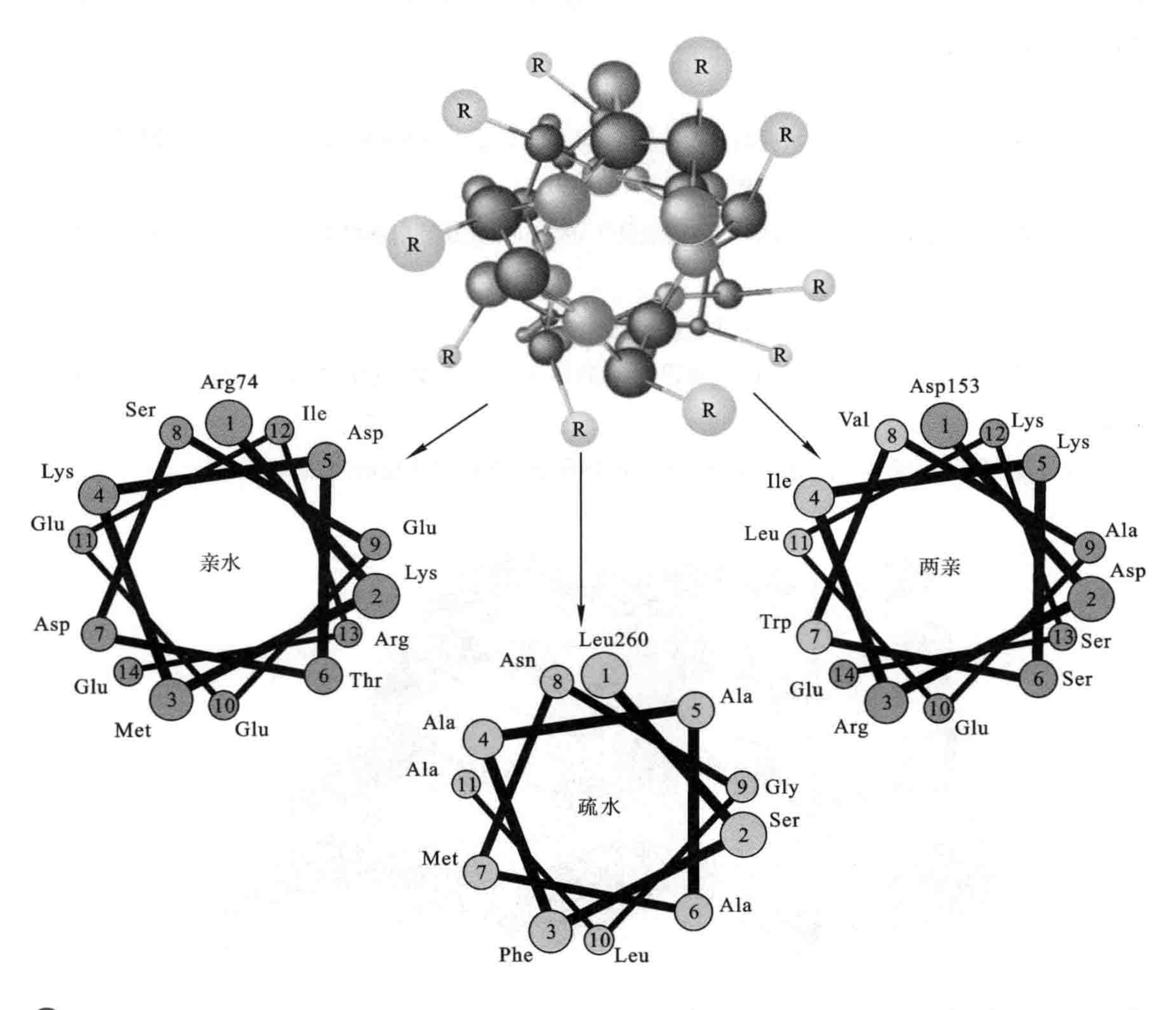


图 2-13 两亲螺旋的螺旋轮作图

除了 α 螺旋以外，Pauling 还提出了 3_{10} 螺旋、2.2_7 螺旋和 π 螺旋结构（图 2-14）。此外在胶原蛋白分子中还有一种叫三股螺旋的二级结构(参见胶原蛋白)。其中 3_{10} 螺旋是 α 右手螺旋的过渡形式，其瘦而长，主链 n 位氨基酸残基的 C═O 与 n+3 位残基的 N—H 形成氢键，每圈只有 3 个氨基酸残基，被氢键封闭的环含有 10 个原子；2.2_7 螺旋比较紧，每圈只有 2.2 个氨基酸残基，其主链 n 位氨基酸残基的 C═O 与 n+2 位残基的 N—H 形成氢键，氢键环共包括 7 个原子；π 螺旋也称为 4.4_{16} 螺旋，每圈有 4.4 个残基，其主链 n 位氨基酸残基的 C═O 与 n+5 位残基的 N—H 形成氢键，氢键环共包括 16 个原子。

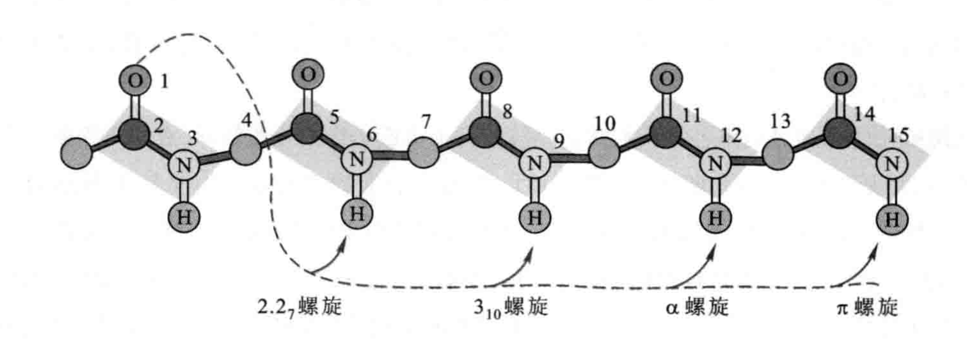


图 2-14 4 种螺旋的氢键形成

以上几种螺旋都可以在体外通过模型建立得到，但存在于蛋白质的螺旋主要还是 α 螺旋。有人在蛋白质数据库中搜寻过 3 万个左右的螺旋片段以后，发现它们几乎都是右手 α 螺旋，只有一例 π 螺旋、一例左手 α 螺旋和一例自由的 3_{10} 螺旋，并且在一些长的 α 螺旋的两端发现有很短的 3_{10} 螺旋，还没有发现 2.2_7 螺旋。

（二）β 折叠

β 折叠（β-sheet）又称为 β 折叠片层（β-plated sheet），这是 Pauling 和 Corey 继发现 α 螺旋结构后，同年发现的又一种重要的蛋白质二级结构。

与 α 螺旋相比，β 折叠是肽链的一种更加伸展的结构，主链呈扇面状展开（图 2-15），其主要特征包括：

（1）肽段几乎完全伸展，肽平面之间呈锯齿状。

（2）相邻肽段呈现平行排列，相邻肽段之间的肽键形成氢键，其中的每一股肽段称为 β 股（β strand）。

（3）R 基团垂直于相邻两个肽平面的交线，并交替分布在折叠片层的两侧。

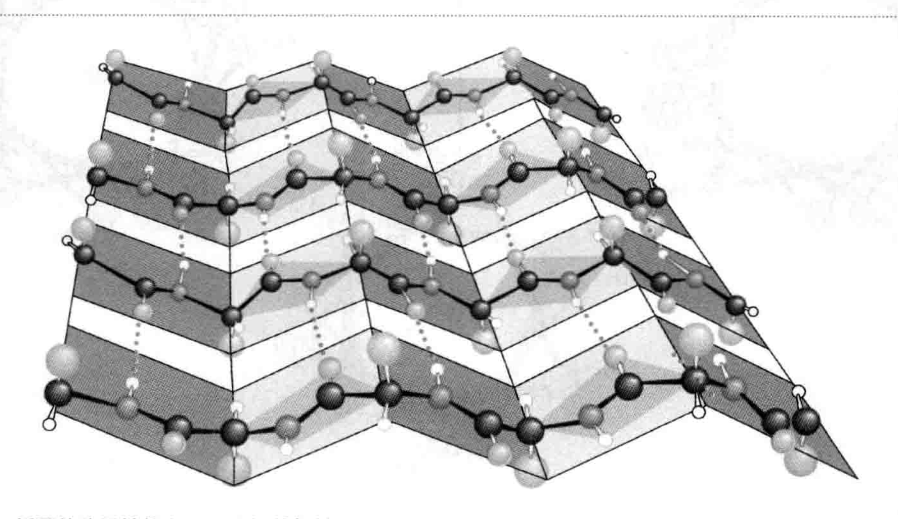


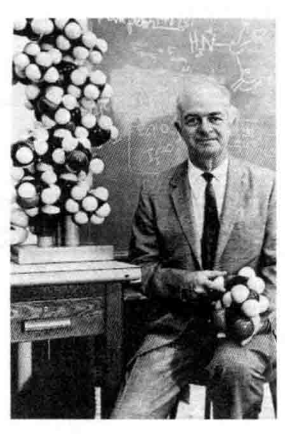
图 2-15 β 折叠的片层结构和 β 股之间的氢键

框 2-2 生化大师的传奇——感冒期间的感悟

在 20 世纪 30 年代,William Astbury 发现潮湿的羊毛或发丝一旦发生显著的伸展,其 X 射线纤维衍射显示出的结构就有剧烈的变化。他所得到的数据似乎表明,没有伸展的羊毛和毛发纤维形成的是一种螺旋状的结构,其螺距约为 0.51 nm。

Astbury 起初提出的是一种弯曲的链结构,后来又提出没有弯曲的纤维形成的是螺旋(他称之为 α 型);而伸展可导致螺旋解旋,形成伸展的结构(他称之为 β 型)。

虽然 Astbury 提出的模型细节是错误的,但是其核心内容却是正确的,与后来标准的二级结构——α 螺旋和 β 折叠相对应。事实上,提出 α 螺旋和 β 折叠的三位科学家 Linus Pauling、Robert Corey 和 Herman Branson 保留了 Astbury 的命名规则。Hans Neurath 首先显示了 Astbury 模型的细节是错误的,原因是模型导致原子在空间上发生冲突。Neurath 的相关论文和 Astbury 得到的数据也启发了 H. S. Taylor、Maurice Huggins 和 Bragg 等人,为了解释角蛋白的结构尝试用新的模型去取代 Astbury 的模型。



Linus Pauling

在现代 α 螺旋模型建立的过程中,有两个关键的进展十分重要:一是对氨基酸和肽的晶体结构的确定,以及 Pauling 对肽键平面性质的正确预测,使得对涉及的化学键的几何形状得以正确处理;二是敢于放弃每一圈螺旋的残基数目必须是整数的观念。两个进展的融合发生在 1948 年的早春,Pauling 因感冒住院。无聊至极,Pauling 在思考蛋白质的结构时,用铅笔在白纸上画出大致正确的多肽链结构,并尝试将纸折叠成螺旋状,同时小心地让肽平面结构不被破坏。凭借对化学键特别是氢键结构的深入了解,在经过数次尝试以后,他提出了靠氢键来稳定的螺旋结构。在出院以后,Pauling 和他的同事 Corey、Branson 一起对模型进一步优化,最后在 Pauling 生日的那一天(1951 年 2 月 28 日)将相关论文投给美国科学院院刊(*PNAS*)并很快得以发表。事实证明,Pauling 提出的 α 螺旋模型正确得难以置信,包括键长这个细节(其精度在之后 40 多年内都没有受到挑战)。但他当时并没有考虑螺旋的手性,事实上,发表在 *PNAS* 上的那篇题为"The Structure of Proteins: Two Hydrogen-Bonded Helical Configurations of the Polypeptide Chain"的论文中画的 α 螺旋结构的示意图是左手螺旋。现在我们知道,自然界的 α 螺旋几乎都是右手螺旋。

α 螺旋的发现是生物化学发展史上的里程碑事件,特别是 Pauling 在提出 α 螺旋的过程中所使用的体外模型建立的方法,后来被 Watson 和 Crick 所采用。现在生物化学和分子生物学研究之中经常采用的计算机辅助分子建模,实际上就是 Pauling 最初使用的体外模型建立的延伸。

Pauling 因 α 螺旋的发现而荣获 1954 年的诺贝尔化学奖,后来他还获得过 1962 年的诺贝尔和平奖。在诺贝尔奖历史上,Pauling 至今仍然是唯一一位独享两次诺贝尔奖的人。

(4) 肽段的走向有正平行和反平行两种。正平行经常简称为平行,指相邻 β 股的 N 端位于同侧,反平行正好相反。在反平行折叠中,氢键的 3 个原子(N—H—O)几乎位于同一直线上,因此反平行折叠更加稳定(图 2-16),其存在的概率就更大。

(5) 反平行 β 折叠的每一个氨基酸残基上升 0.347 nm,二面角(ϕ,ψ)约为(-135°,140°);平行 β 折叠的每一个氨基酸残基上升 0.325 nm,二面角(ϕ,ψ)约为(-120°,105°)。

因为平行 β 折叠的稳定性低于反平行 β 折叠,所以平行 β 折叠一般含有较多的 β 股,很少低于 5 个,并且 β 股总是被包埋在蛋白质的内部;而反平行 β 折叠含有的 β 股可以低到 2 个,其一面经常暴露在水相之中。

不管是哪一类 β 折叠,在两边的 β 股上总会有一些 C═O 和 N—H 没有参与形成 β 折叠内部的氢键,因此要通过其他的方式满足它们形成氢键的"欲望"。这些方式包括:①与水分子形成氢键,如果 β 股正好暴露在水相之中;②与其他二级结构(如 α 螺旋)上的亲水氨基酸残基的侧链形成氢键,如果 β 股在蛋白质的内部;③如果 β 股在蛋白质的表面,就与其他有相同问题的蛋白质分子的表面 β 股形成 β 折叠;④构成 β 折叠的所有 β 股在空间上会发生一定的扭曲,以便让第一股和最后一股靠

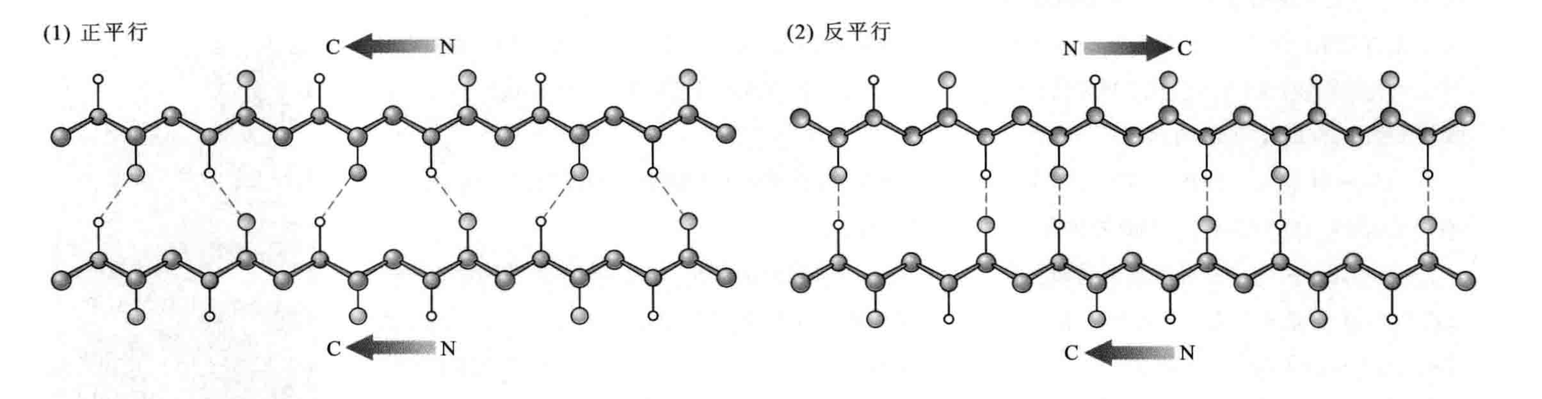
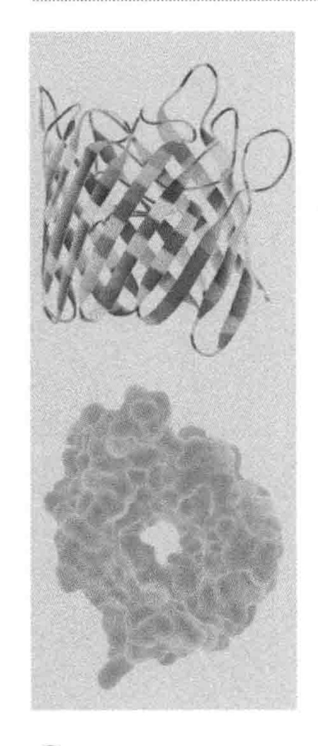


图 2-16　平行 β 折叠和反平行 β 折叠的结构比较

近，即首尾相连而形成氢键，以致这些 β 股聚集在一起，形成一种伸展的右手扭曲结构（extended right-handed twists）。这样的结构实际上是一种结构模体（见后），多存在于许多蛋白质分子的核心。

有时，两个蛋白质在特定的条件下，可以各提供一个 β 股形成 β 折叠，从而导致两者瞬间结合，形成临时的复合物（例如真核细胞内的 Raf 蛋白和 Rap 蛋白）。这种情况经常出现在细胞的信号转导过程中。

如果 β 股交替出现疏水残基和亲水残基，那么可以形成两亲 β 折叠。这样的结构常见于孔蛋白（porin）。孔蛋白是一种膜蛋白，由两亲 β 折叠形成的 β 桶结构（β-barrel）构成极性分子进出细胞的孔洞。在 β 桶结构中，非极性残基的侧链伸向膜脂的疏水区，极性残基在内部形成通道，允许极性分子通过膜（图 2-17）。


图 2-17　孔蛋白中由两亲 β 折叠构成的 β 桶结构

相比于 α 螺旋，参与形成 β 折叠的 β 股在空间上处于更加伸展的状态，因而那些在 α 螺旋中容易产生空间位阻的氨基酸残基（在 β 碳上有分支的）能在 β 折叠中找到合适的位置，即侧链基团庞大的氨基酸残基更倾向于形成 β 折叠。当然，那些在 β 碳上无分支的氨基酸残基也适合存在于 β 股上，但大的疏水侧链更容易紧密地结合在一起，故它们出现在 β 折叠中的频率更高。这些氨基酸残基有 Val、Ile、Phe、Tyr、Trp 和 Thr。Pro 的刚性环结构限制了它的 ϕ 值（−85°~−35°），而 β 折叠的 ϕ 值在 −119° 左右，所以 Pro 从来不会出现在 β 折叠中。

Quiz7　学到现在，你知道 Pro 有哪些性质与其他氨基酸不一样？

（三）β 转角

为了在一条肽链上形成 β 折叠，或将一条多肽链上的 α 螺旋和 / 或 β 折叠组合在一起，通常需要多肽链的主链发生弯曲或回折，而 β 转角结构就可以达到这样的目的。

β 转角也称为 β 弯曲（beta-bend）、β 回折（beta-reverse turn）、紧密转角（tight turn）或发夹结构（hairpin structure），这种结构是指伸展的肽链形成 180° 的 U 形回折。其主要特征包括：

（1）主链以 180° 的回折而改变了方向。

（2）由肽链上 4 个连续的氨基酸残基组成，其中 n 位氨基酸残基的 C═O 与 n+3 位氨基酸残基的 N—H 形成氢键。

（3）Gly 和 Pro 经常出现在这种结构之中。这是因为 Gly 的 R 基团最小，很容易调整其在 β 转角中的位置，降低与其他残基（尤其是 R 基团大的残基）之间可能形成的空间位阻，而 Pro 则具有相对刚性的环状结构和固定的 ϕ，在某种程度上能迫使转角的形成。

（4）有利于反平行 β 折叠的形成。这是因为 β 转角改变了肽链的走向，促进相邻的肽段各自作为 β 股，形成 β 折叠。

β 转角有三种类型，前两种类型（Ⅰ型和Ⅱ型）的差别仅仅在于第二个残基和第三个残基形成的肽平面的朝向（相差 180°）（图 2-18），第三种类型（Ⅲ型）则有重复的二面角，(ϕ,ψ) 等于（−60°，−30°），

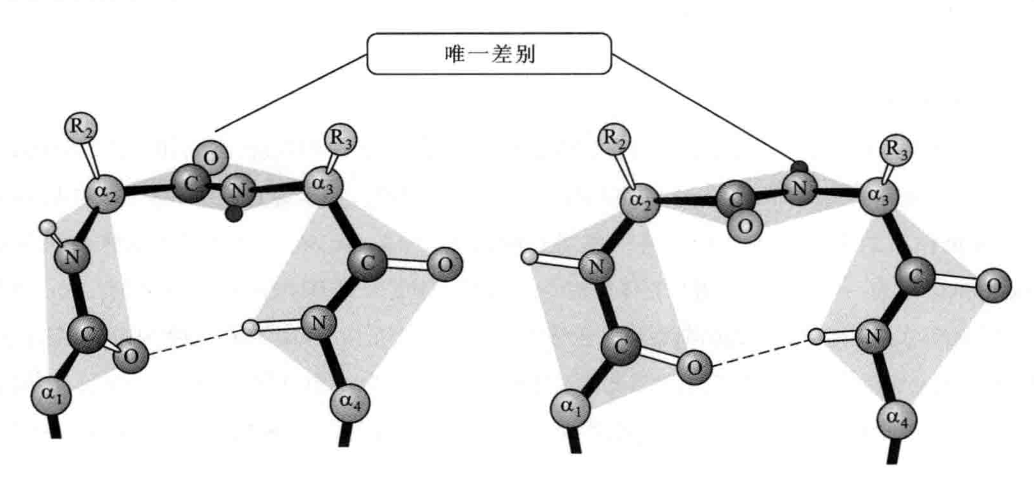


图 2-18 两类 β 转角的结构

与 3_{10} 螺旋接近。Ⅰ型转角更为常见，它的 $n+2$ 位氨基酸残基最适合 Pro；而Ⅱ型的 $n+1$ 位氨基酸残基最适合 Pro，$n+2$ 位适合 Gly 和其他小的亲水氨基酸。无论是哪一种类型，$n+3$ 氨基酸残基最适合 Gly，而 n 位氨基酸残基经常是 Asn、Asp、Ser 和 Cys。

在构成 β 转角的 4 个氨基酸残基之中，只有 n 位氨基酸残基的 C=O 与 $n+3$ 位残基的 N—H 形成氢键，其余位置的氢键供体和受体之间并没有相互形成氢键。为了让它们也能参与形成氢键，需要将它们放在蛋白质的表面，以方便与水分子共筑氢键。正因为如此，β 转角通常出现在球状蛋白质的表面。已发现蛋白质的抗体识别、磷酸化、糖基化和羟基化位点经常出现在 β 转角和紧靠 β 转角的位置。此外，在水溶液中的小肽分子上也发现类似 β 转角的结构。

由于 β 转角能够改变肽链的走向，因此它对球状蛋白质的形成十分重要，但对于纤维状蛋白质来说，它的肽链在空间上只有一个取向，故一般没有 β 转角。

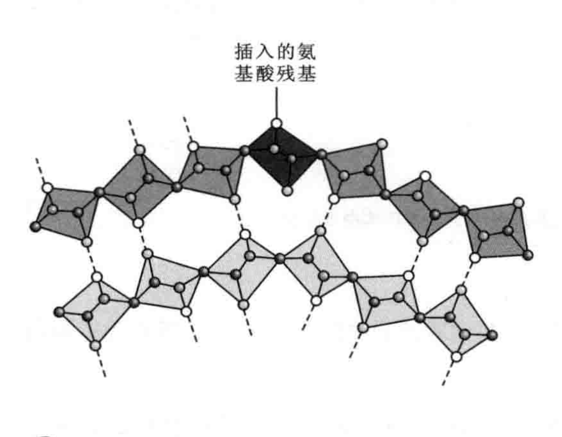


图 2-19 β 凸起

（四）β 凸起

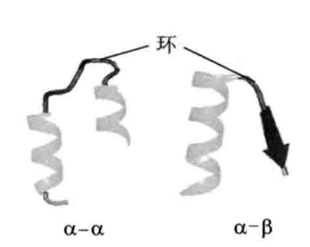
如图 2-19 所示，当构成 β 折叠的两个 β 股其中的一股多出一个氨基酸残基的时候，原来连续的氢键结构即被打破，这个多出的氨基酸残基由于没有能与之形成氢键的供体或受体，只能突出在外，从而使肽链发生轻微的弯曲，形成了 β 凸起。β 凸起主要存在于反平行 β 折叠之中，只有约 5% 的 β 凸起出现在平行的 β 折叠结构之中。β 凸起也能改变多肽链的走向，但没有 β 转角那样明显。

Quiz8 为什么 β 凸起更容易出现在反平行 β 折叠之中？

（五）无规则卷曲与环

在蛋白质分子中，除了上述几种有规则的二级结构以外，还有一些极不规则的二级结构，它们柔性强，涉及的二面角（ϕ，ψ）变化大，但也不是任意变动的，这些结构统称为无规则卷曲。

环在很长一段时间里被归为无规则卷曲的一种，但随着对其结构和功能认识的深入，人们越来越倾向于将其视为一种独立的二级结构形式。环作为蛋白质三维结构中最重要的动态结构元件，一般位于球状蛋白质的表面，在很多地方充当将有规则的二级结构联系在一起的纽带（图 2-20），其侧链和主链部分通常含有各种结合位点和功能位点，并以运动的方式作为控制与蛋白质相互作用的配体进入的“门槛”。已发现的与环有关的生物功能包括：分子识别、与其他蛋白质相互作用、与配体结合、参与酶的催化或控制酶的催化等。从进化的角度来看，相同的蛋白质家族的不同成员在功能上的差异通常是由蛋白质表面的结构变化引起的，而这些变化经常发生在环上。因此，环结构的突变是蛋白质



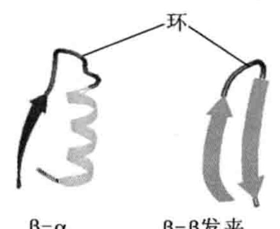


图 2-20 存在于有规则的二级结构之中的各种环

进化产生新活性和新功能的常见手段。

三、蛋白质的三级结构

三级结构是指构成蛋白质的多肽链在二级结构的基础上，进一步盘绕、卷曲和折叠，形成的特定空间结构，它包括了肽链上所有原子的空间排布。三级结构通常由模体（motif）和结构域（domain）组成。一种蛋白质的全部三维结构又可以称为它的构象（conformation）。但注意不要将构象与构型（configuration）混为一谈。构型是指在立体异构中，一组特定的原子或基团在空间上的几何布局。两种不同构型的转变总是伴随着共价键的断裂和重新形成。然而，一个蛋白质可以存在几种不同的构象，但构象的转变仅仅是单键的自由旋转造成的，没有共价键的断裂和形成。构成蛋白质的多肽链上存在多个单键，因此一种蛋白质在理论上可能具有许多不同的构象。然而，在生理条件下，一种蛋白质只会采取一种或几种在能量上有利的构象。

Quiz9 肽键从反式变成顺式是构象还是构型的变化？

（一）稳定三级结构的化学键

稳定三级结构的化学键主要是次级键，包括氢键、疏水键、离子键、范德华力（图 2-21）。有的金属蛋白还借助于金属配位键来稳定它们的三级结构。此外，属于共价键的二硫键也参与稳定许多蛋白质的三维结构。

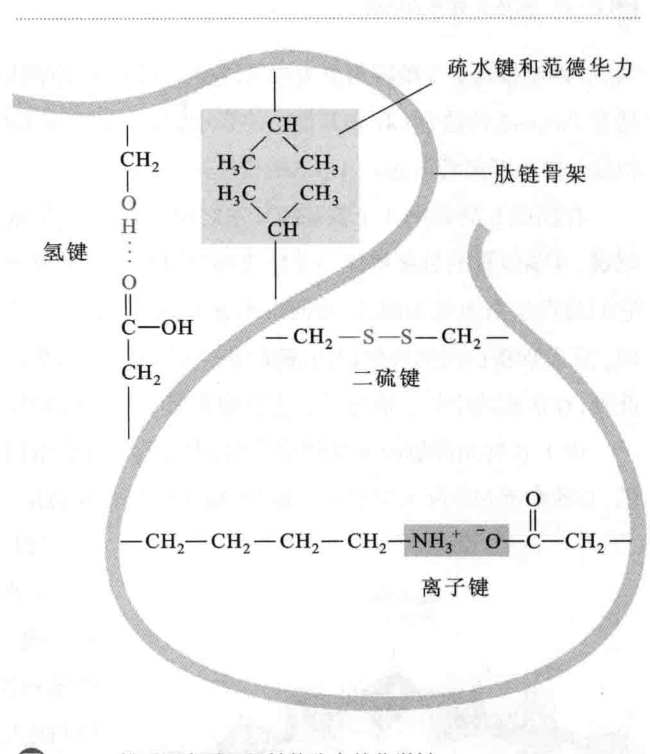


图 2-21 维系蛋白质三级结构稳定的化学键

（1）氢键 凡是与电负性很强的原子（如 O、N 或 F）相连的氢原子（氢键供体），由于带部分正电荷，可以和另一电负性较强的带部分负电荷的原子（氢键受体），通过静电吸引相连接，以这种方式形成的键叫氢键。氢键除了用来稳定蛋白质的二级结构以外，在维持三级结构方面也起着重要作用。不过，稳定二级结构的氢键供体和受体来自主链，而稳定三级结构的氢键供体和受体一般则来自 R 基团，例如羟基氨基酸的—OH 和酸性氨基酸的—COOH。

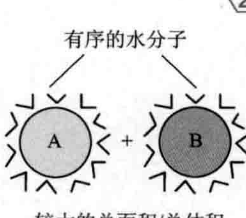
氢键的键能为 12~30 kJ/mol，比共价键弱得多。但由于蛋白质分子中存在许多氢键，其在维持蛋白质三级结构的稳定性中仍有很大贡献。

（2）离子键 离子键是带相反电荷基团之间的静电引力（electrostatic interactions），其键能为 20 kJ/mol，在蛋白质中常称为盐键（salt bond）或盐桥（salt bridge）。在生理 pH 下，肽链上的碱性氨基酸残基侧链带正电，酸性氨基酸残基侧链带负电，此外游离的 N 端氨基和 C 端羧基也分别带正、负电荷，所有这些带电荷的基团都可以参与形成盐键。

（3）疏水键 构成蛋白质的疏水残基上的各种非极性 R 基团避开水相，互相聚集在一起而形成的作用力称为疏水键（hydrophobic bond），也称为疏水作用力，其键能 <40 kJ/mol。在水溶液中，疏水基团会躲避周围的极性溶剂而自发地聚集在一起，从而在分子内部形成疏水核心。这种疏水相互作用在维持蛋白质的三级结构中起着十分重要的作用。如果一个多肽的所有氨基酸残基都是亲水的，那么用来驱动三级结构形成的力将十分有限。而如果一个多肽既含有亲水氨基酸又含有疏水氨基酸，则有利于蛋白质在溶液中快速地折叠，并到达最终的构象状态。

需要特别注意的是，疏水键并不是发生在疏水基团之间的一种主动吸引力，而是一种受热力学第

二定律驱动的作用力。将疏水基团集中在蛋白质内部的疏水核心在能量上是有利的，因为它降低了疏水基团与水分子之间不利的作用，反而使疏水基团之间的范德华力大大增加。如图 2–22 所示，如果两个疏水基团（A 和 B）单独存在的话，其表面会附着更多有序的水分子，所以熵比较小；然而，如果两个疏水基团聚合在一起（A–B）的话，由于球面积是半径的平方，而体积是半径的立方，那么将有一部分水分子从疏水基团的表面释放出来，进入无序状态，这会导致系统的熵增加。另外，当两个疏水基团因为疏水作用力而靠近，这也为它们之间建立范德华力创造了条件。正因为如此，疏水键和范德华力往往相伴相随，一起对稳定蛋白质的三维结构、核酸的双螺旋结构以及生物膜的脂双层结构起决定性作用。



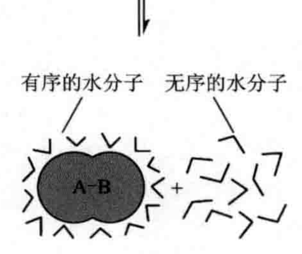


图 2–22　疏水键的化学本质

（4）范德华力（van der Waals force）　范德华力是在两个相邻的不带电荷的非成键原子（non-bonded atom）之间的作用力，这种作用力可能是吸引力也可能是排斥力。吸引性的范德华力源于相邻的不带电荷的非成键原子因为电荷密度的波动，引发的诱导偶极（induced dipole）之间的相互吸引；排斥性范德华力产生的原因是相邻的不带电荷的非成键原子靠得太近，但并没有诱导偶极，反而因电子云的重叠引发电子与电子之间的排斥。

尽管范德华力非常弱，键能只有 0.4~4 kJ/mol，但在蛋白质分子中，这样的作用力大量存在，因此其对蛋白质三级结构的形成和稳定所起的作用不容小觑。

（5）配位键　配位键是一种特殊形式的共价键，它对某些金属蛋白（metalloprotein）的三级结构具有一定的稳定作用。以羧肽酶 A（carboxypeptidase A）为例，这种由 307 个氨基酸残基组成的蛋白质大致折叠成球形，这和 1 个 Zn^{2+} 与其 His 69、Glu 72、His 196 的侧链以及 1 个水分子配位结合形成四面体结构有一定关系。

（6）二硫键　二硫键也称为二硫桥，它是在两个半胱氨酸残基的 S 原子之间形成的共价交联。含有二硫键的蛋白质一般是分泌蛋白或者细胞膜蛋白（如胰岛素、胰岛素的受体、动物消化道内的各种蛋白酶），而胞内蛋白很少含有二硫键。其原因在于：一是胞外的环境比较恶劣（如胃液的 pH 在 2 以下）；二是蛋白质也比较小，故疏水作用力相对要弱；三是胞外环境的氧化性较强。这三种因素使蛋白质必须借助键能最强的二硫键来加固它们的三级结构。

（二）研究三级结构的方法

确定蛋白质三级结构的方法主要有两种：一种是 X 射线晶体衍射，另一种是核磁共振影像（nuclear magnetic resonance imaging，NMR）。这两种方法各有千秋，有时也用来研究蛋白质的二级结构。前一种方法首先需要得到蛋白质的晶体，只有在得到晶体以后，才可以对蛋白质进行 X 射线衍射分析。但是，想得到一种蛋白质的晶体并非易事，特别是膜蛋白，某些蛋白质的晶体也许永远都得不到。此外，X 射线衍射分析得到的并不是原子的直接图谱，而是电子密度图谱。因此需要利用专业的傅里叶逆变换（inverse Fourier transform）去分析电子密度，再还原成三维结构（图 2–23）。该方法的缺点是：得到的是蛋白质分子在晶体状态下的三维结构。这种结构与蛋白质分子在细胞内的天然结构可能有较大的差别，原因是晶体中的蛋白质分子相互间是紧密而有规律地排列在一起的，运动性较

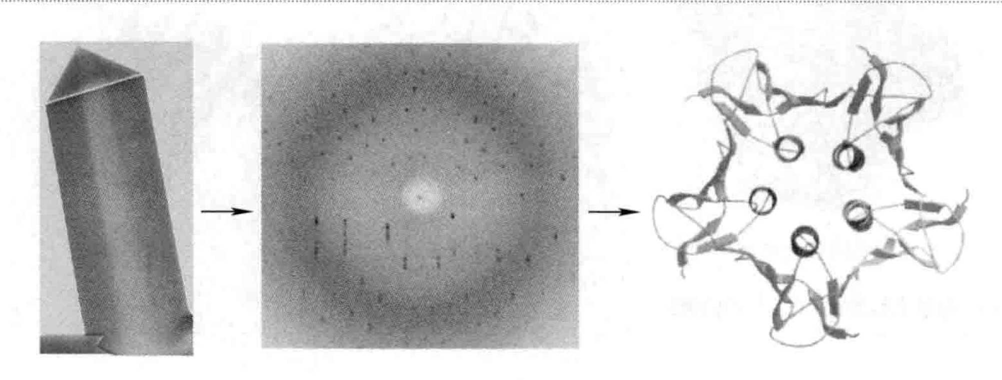


图 2–23　X 射线晶体衍射的电子密度图和被还原出来的三维结构

差，而细胞中的蛋白质分子大多数处于溶液状态，周围是水分子和其他生物分子，运动性较强。而且，有些蛋白质只能稳定地处于溶解状态，无法结晶。后一种方法是根据原子核在磁场下的振动情况来分析蛋白质结构，由瑞士化学家 Kurt Wüthrich 建立。它可以直接对在溶液中的蛋白质结构进行研究，得到的三维结构也更接近蛋白质在生理条件下的真实构象。此外，通过改变溶液的性质，可以模拟出细胞内的各种生理条件，以观察周围环境的变化对蛋白质分子构象的影响。但此种方法限于研究由少于 120 个氨基酸残基组成的蛋白质，这就大大地限制了它的使用。而 X 射线晶体衍射的方法对蛋白质分子的大小没有限制。

自 1985 年 NMR 建立以来，已有几百种蛋白质的空间结构通过这种方法获得，约占已知空间结构的蛋白质总数的 20%。其中有些蛋白质也利用 X 射线晶体衍射测定过。比较两种方法得到的结果，可以看出大体上是相同的，但局部的表面区域由于它们所处的环境不一样而呈现出明显的差异。另外，许多蛋白质在晶体状态下仍然具有在细胞生理状态下的生物活性（如与配体结合、催化反应等）。由此可见，两种方法得到的结果都是可靠的。

（三）三级结构的表示方法

在得到一种蛋白质的三维结构以后，可以用不同的模型将其展示出来。几种常见的模型包括（图 2-24）：①线框（wire-frame）模型，用线框表示蛋白质主链及其侧链，以显示蛋白质分子在三维空间上的轮廓；②球棍（ball and stick）模型，用“球”代表成键原子，“棍”代表原子之间的共价键，以此显示蛋白质分子的三维结构；③棍式（stick）模型，用“细棍”代表原子之间的共价键，但原子本身并没有显示出来；④空间填充（space-filling）模型，用重叠的球来表示成键原子，球的直径由范德华半径决定，此模型能更真实地显示出蛋白质分子的外形，但很难从模型中看见它的键角，很多重要的结构信息也没有表现出来；⑤骨架（backbone）模型，用粗的线条显示蛋白质主链在空间上的走向；⑥丝带（ribbon）模型，用不同形状的丝带显示构成蛋白质分子的各种二级结构在三维空间内的排布，此模型虽不能显示构成蛋白质分子的各个原子在空间上的具体排布，但能显示与蛋白质三维结构相关的重要的二级结构信息。

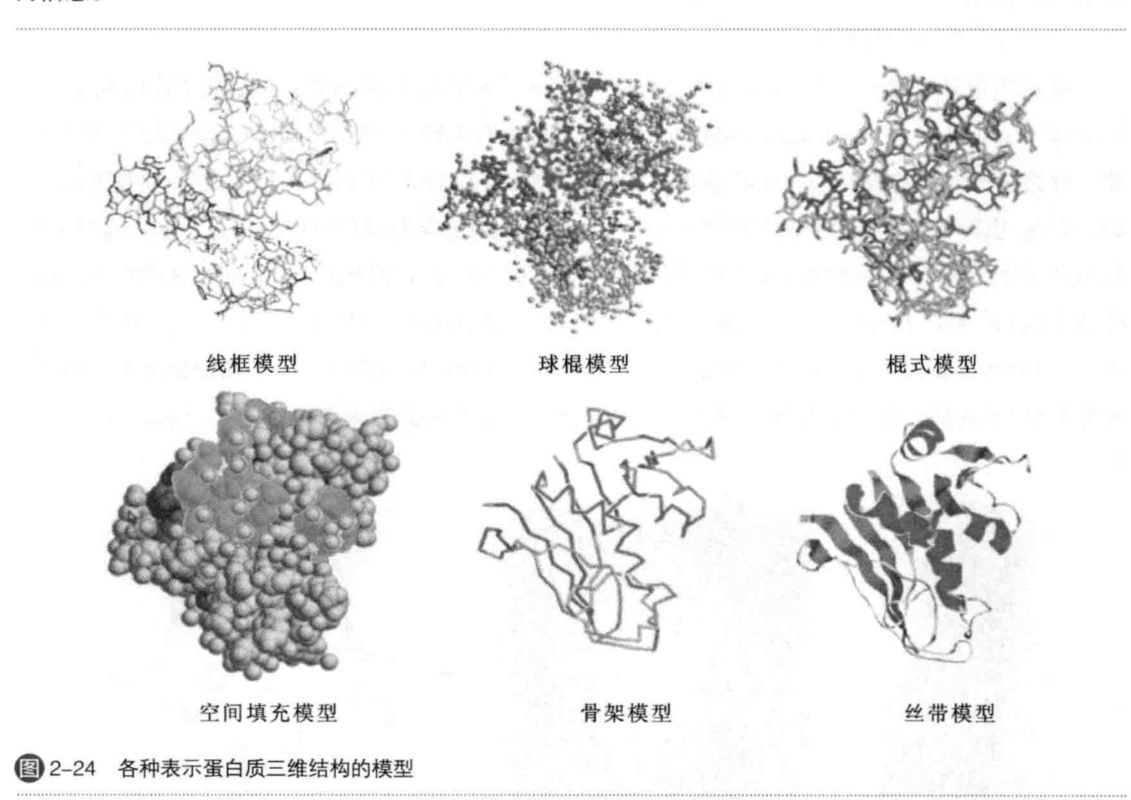


图 2-24　各种表示蛋白质三维结构的模型

(四) 三级结构的结构部件

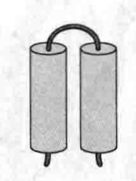
1. 模体

在结构生物学(structural biology)上,模体这个概念有两种不同的用法。第一种用在一级结构上,特指具有特殊生物化学功能的特定氨基酸序列或DNA分子上的碱基序列,被称为序列模体(sequence motif)。例如,很多DNA结合蛋白(DNA binding protein)的一级结构含有一段$CX_{2-4}CX_{12}HX_3H$序列(X代表任何一种氨基酸残基),它是形成锌指结构(zinc finger)的特征序列(参看第四十三章"真核生物的基因表达调控")。再如,能够与整联蛋白(integrin)结合的细胞外基质蛋白一般含有RGD序列模体。还有在细胞质核糖体上合成的、最后定位并留在内质网腔的蛋白质都含有KDEL序列模体(参看第四十章"蛋白质的翻译后加工及其定向和分拣")。模体的第二种用法表示的是结构模体(structural motif),相当于超二级结构(super-secondary structure)。如果一种结构模体对应于一种特定的生物功能,那么这种结构模体可称为功能模体(functional motif)。

结构模体作为结构域的组分,介于蛋白质二级结构和三级结构之间,由相邻的二级结构单元彼此相互作用,组合在一起,排列成规则的、在空间结构上易于辨认的二级结构组合体,并充当三级结构的构件。多数情况下只有疏水残基的侧链参与这些相互作用,而亲水侧链多在分子的外表面。常见的结构模体包括(图2-25):①卷曲螺旋(coiled coil);②螺旋-环-螺旋(helix-loop-helix,HLH);③螺旋-转角-螺旋(helix-turn-helix,HTH);④ Rossmann折叠(Rossmann fold);⑤希腊钥匙模体(the Greek key motif);⑥β发夹(β-hairpin)。



β发夹



螺旋-转角-螺旋

图2-25 几种结构模体

卷曲螺旋是一种常见的α螺旋聚合体,一般是由两股或三股α螺旋组装而成的一种超螺旋结构(图2-26),每一股螺旋一般含有HXXHCXC七肽重复序列(heptad repeat),这里的H表示疏水氨基酸残基(经常是Leu),C表示带电荷的氨基酸残基,X表示其他类型的氨基酸残基。这种有规律的七肽重复序列使螺旋之间能够通过螺旋表面的疏水补丁聚合在一起,形成二聚体或多聚体。卷曲螺旋不仅作为三级结构的模体参与蛋白质的折叠,而且还参与蛋白质与蛋白质之间的相互作用,因为蛋白质之间可以通过它形成二聚体或多聚体复合物,例如α-角蛋白之间就可以通过这种方式形成二聚体(参看第三章"蛋白质的功能及其与结构之间的关系")。

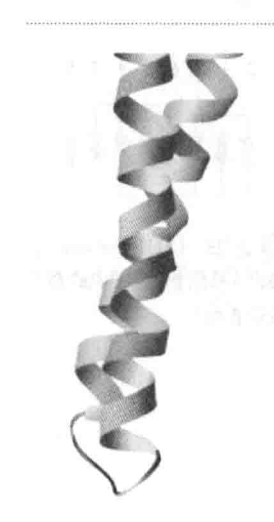


图2-26 双股卷曲螺旋

HLH又称为EF手相(EF-hand),是一种典型的功能模体。它是由E螺旋、F螺旋和螺旋之间的一个环组成(图2-27),已被发现存在于多种与Ca^{2+}结合的Ca^{2+}传感器蛋白上(参看第十七章"激素及其受体介导的信号转导"),Ca^{2+}在环上与蛋白质结合。螺旋-环-螺旋里面有一类叫碱性螺旋-环-螺旋(basic helix-loop-helix,bHLH),其中的一股螺旋含有碱性氨基酸残基。许多参与基因转录的转录因子(如c-Myc和MyoD)含有这种模体结构。这些转录因子通过碱性螺旋上的亲水氨基酸残基,识别DNA双螺旋大沟内特定的碱基序列,并形成氢键,从而与DNA分子上的特定序列位点结合(参看第四十三章"真核生物的基因表达调控")。

Quiz10 为什么将这两段螺旋称为E螺旋和F螺旋?

Quiz11 你认为α螺旋上哪些氨基酸残基最容易与DNA上的碱基产生特异性的相互作用?

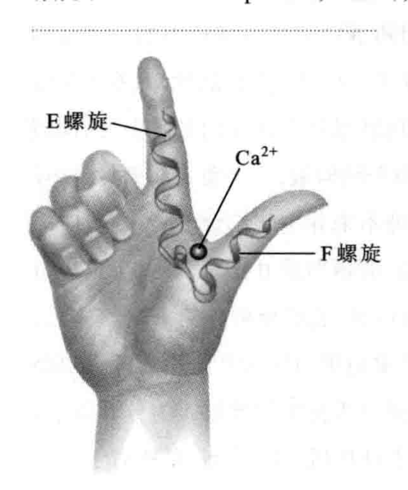
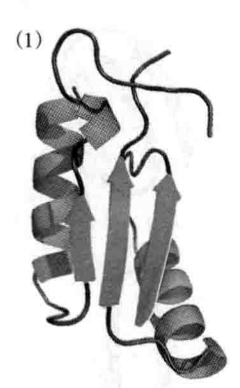
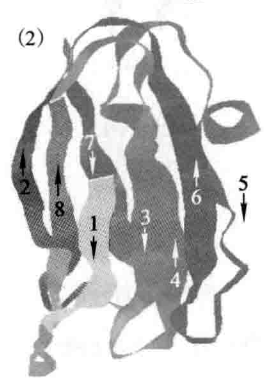


图2-27 EF手相结构

HTH模体是另外一种重要的功能模体,它也出现在许多与DNA特异性结合的蛋白质分子上(如Cro、CAP和λ阻遏蛋白,参看第四十二章"原核生物的基因表达调控"),其中与DNA结合的也是两股α螺旋中的一股。这股螺旋主要也是通过表面亲水氨基酸残基的侧链,与DNA分子上的特殊碱基序列形成氢键而结合。

Rossmann折叠由两个相连的βαβ单位组成,每一个βαβ单位由两段平行的β股和一段作为连接链的α螺旋组成,进而形成βαβαβ结构。同时在紧密结合的折叠和螺旋之间经常有一个疏水的核心[图2-28(1)]。这种模体能结合辅酶Ⅰ和辅酶Ⅱ,故存在于许多以NAD^+或$NADP^+$为辅酶的脱氢酶(如乙醇脱氢酶)中。





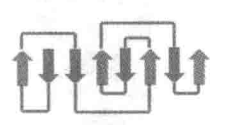
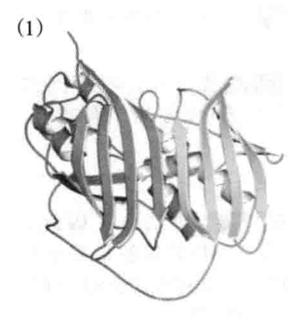


图 2-28 (1) Rossmann 折叠;(2) 希腊钥匙(图中的箭头表示 β 股)



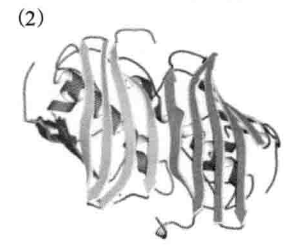


图 2-29 两个相邻结构域的完美重叠(Petsko 等,2004)

希腊钥匙模体是一种全 β 折叠聚合体,存在于许多不同类型的蛋白质中,因在拓扑学上像古代花瓶上的希腊钥匙而得名[图 2-28(2)]。例如,白蛋白原(pre-albumin)和质体蓝素(plastocyanin,PC)就含有这种模体。

β 发夹因形似发夹而得名,它是蛋白质所有模体中最简单的一种,由两段反平行的 β 股和一段连接小环(发夹环)组成。β 股之间形成氢键,小环含有 2~5 个氨基酸残基。这种模体结构有时会出现在一些小肽分子上。

2. 结构域

某些蛋白质(如胶原蛋白)呈纤维状,其肽链向同一个方向伸长,但多数蛋白质会卷曲成致密的球状结构。M_r<2 × 10^4 的蛋白质通常呈一个简单的球形(平均直径为 2~3 nm),而较大的蛋白质一般会折叠成两个或多个相对独立的球状区域。这些相对独立的球状结构和 / 或功能模块称为结构域(domain)。每一个结构域通常是独自折叠形成的,内部都有一个疏水的核心,并含有一个或几个模体结构。疏水的核心是结构域稳定存在所必需的。

结构域大小有别,但通常不会大于最大的单结构域蛋白(大约由 250 个氨基酸残基组成),其中大多数结构域由 200 个以下氨基酸残基组成,由 51~150 个氨基酸残基构成的结构域约占 49%。迄今发现的最大的结构域含有 907 个氨基酸残基。大的结构域疏水核心较大,所以稳定性要好。小的结构域疏水核心较小,因而稳定性要差,经常需要借助于金属离子的结合或二硫键来加固。

不同的蛋白质含有的结构域数目也可能不一样,以 2~3 个较为普遍,其中有一个蛋白质被发现含有多达 13 个结构域。另外,某些结构域在同一个蛋白质分子上有多个拷贝,即一个蛋白质分子含有多个相同或相似的结构域。有时,两个相邻的结构域之间除了一些起连接作用的外部环和小的肽段以外,可以发生几乎完美的重叠(图 2-29)。一般认为,这样的蛋白质是通过基因重复(gene duplication)的方式进化而来的,即一个结构域是另外一个结构域的复制品。对于某些蛋白质,复制的结构域与母体结构域在序列上具有某种一致性,而对于其他一些蛋白质,经过多年的进化,序列跟母体相比已变得面目全非。复制的结构域是否与原来的结构域具有一致性,则取决于复制事件已经历了多长时间以及指导它趋异的功能对原序列的影响。显然,基因复制的越古老,积累的突变就越多,从而使序列之间的亲缘关系变得模糊。以大肠杆菌的硫酯酶(thioesterase)为例,它有两个结构域在大小和形状上几乎完全相同,但氨基酸序列差之甚远,这显然是几百万年的随机突变造成的。

结构域在结构上是相对独立的,在功能上也是如此。许多蛋白质的结构域在特殊的条件下被分开以后,每一个结构域仍然保留各自原有的功能。例如,大肠杆菌 DNA 聚合酶 I 含有 3 个结构域,分别有 DNA 聚合酶、3′- 外切核酸酶和 5′- 外切核酸酶的活性。使用胰蛋白酶可以将其中的一个结构域和另外两个结构域分开。利用这种方法游离出来的小结构域具有 5′- 外切核酸酶活性,另外两个结构域保留 DNA 聚合酶和 3′- 外切核酸酶活性。在许多参与激活基因转录的激活蛋白分子上,结构域所具有的相对独立性更是表现得淋漓尽致。这些激活蛋白一般有两个结构域,一个负责与 DNA 特异性结合,另一个负责激活基因的表达。当将 DNA 结合结构域与一种本来并不能结合 DNA 的蛋白质 A 融合,再将激活基因表达的结构域与另外一种并不能激活基因表达的蛋白质 B 融合以后,令人惊奇的是,如果 AB 之间能够发生特异性相互作用,那么 AB 在相互结合以后,竟然能够作为一个整体,去激活报告基因的表达。有人正是根据这一点,发明了一种研究蛋白质与蛋白质相互作用的酵母双杂交技术(参看第四十四章"重组 DNA 技术")。当然,并不是所有的蛋白质在结构域被人为分开以后,可以保证功能不受到影响。如果一种蛋白质的某项功能由它的两个结构域共同承当,显然对于这样的蛋白质,其结构域的分离必然会影响到它的功能。还有一种情形,一个结构域的功能是调节另外一个结构域的活性,例如某些酶的催化中心在一个结构域上,调节酶活性的别构中心在另外一个结构域上,如果将这两个结构域分开的话,这些酶的活性将不再受原来的与别构中心结合的别构效应物的调节(参看第十二章"酶活性的调节")。

根据占优势的二级结构元件的类型,结构域通常分为四大类:① α 结构域,完全由 α 螺旋组成;② β 结构域,只含有 β 折叠、β 转角和不规则环结构;③ α/β 结构域,由 β 股和起连接作用的 α 螺旋片段组成,其组成单元是 βαβ 模体;④ α+β 结构域,由独立的 α 螺旋区和 β 折叠区组成。第四类与前三类会发生重叠,所以有时难以鉴定。构成这四类结构域的二级结构元件可能有不同的组织方式,而每一种组织就是一种结构模体。此外,还有一些结构域缺乏特定的二级结构元件,需要借助二硫键或金属离子来稳定。

α 结构域经常含有四螺旋束(four-helix bundle)模体,并行使多种生物功能,包括氧气运输、核酸结合和电子传递等。这种模体由 4 个反平行的 α 螺旋聚合而成,每一个 α 螺旋与下一个相交,夹角约为 -20°,于是整个模体向左手弯曲。例如,蚯蚓肌红蛋白(myohemerythrin)(图 2-30)和人生长激素就含有这种模体。另外一种常见的组成 α 结构域的模体是珠蛋白折叠(globin fold),它一般由 8 个左右的 α 螺旋组成,相互间的夹角为 90° 和 50°。这种模体使结构域内部形成一个疏水的口袋,以容纳大的疏水性的血红素辅基。该折叠因在肌红蛋白中首先发现而得名。肌红蛋白只有一个结构域,功能是贮存氧气,构成模体的 8 个 α 螺旋环绕其血红素辅基(图 2-30)。类似的模体还发现在另一类含有血红素辅基但参与电子传递的细胞色素分子上。



蚯蚓肌红蛋白



肌红蛋白

图 2-30 蚯蚓肌红蛋白四螺旋束和人肌红蛋白的珠蛋白折叠(Petsko 等,2004)

由 β 结构域组成的蛋白质包括免疫球蛋白、几种酶(如超氧化物歧化酶)和细胞表面的一些糖蛋白。因为在相邻的 β 股之间没有起连接作用的螺旋,所以全 β 结构域实际上只有反平行的 β 折叠,在 β 股之间起连接作用的是 β 转角和较大的环。



图 2-31 磷酸丙糖异构酶桶的 TIM- 桶结构域

在 β 结构域内的反平行折叠倾向于将它们的亲水一面面向蛋白质的表面,与水环境接触,而将疏水的一面面向疏水的核心,相互间还可以包装在一起,形成 β 桶结构。例如,白蛋白原使用希腊钥匙模体构建 β 桶结构域。

在含有 β 桶的蛋白质中有一亚类,其 β 桶是封闭的。在这种封闭的 β 桶中,可以检测到一圈氢键,即每一个 β 股与左右两个相邻的 β 股形成氢键,这意味着第一个 β 股与最后一个 β 股也形成氢键。在已确定的 β 桶结构之中,β 股的数目从 4 到 20 不等,β 折叠多为反平行或混合型。例如,胰凝乳蛋白酶有两个结构域,每一个结构域都具有反平行的6股β折叠。

α/β 结构域的一个常见的例子是磷酸丙糖异构酶桶(triosephosphate isomerase-barrel,TIM-barrel)结构域或 TIM- 桶折叠(TIM-barrel fold)结构域,它由 8 个重复的 βα 单元组成桶状结构,每一个 α 螺旋位于一个 β 股之后(之间是连接环),桶内侧的 8 个平行 β 股被外侧的 8 个 α 螺旋覆盖,可简写为$(\beta\alpha)_8$(图 2-31)。TIM- 桶折叠是蛋白质数据库(protein data bank,PDB)中最常见的一种结构域,它被发现存在于许多不同的酶家族中。PDB 是一个生物大分子三维结构实验数据库,其网址为 http://www.pdb.org,它由美国布鲁克海文国家实验室(Brookhaven National Laboratory)建立,为人们提供蛋白质信息服务。

四、蛋白质的四级结构

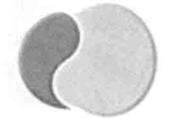
许多蛋白质由一条以上(2~6 条,甚至更多)的肽链组成。它们有可能是由同一种多肽链自组装而成(a_2,一个 a 表示一条肽链),也可能还含有其他不同的肽链(如 ab、a_2b_2 或 abc 等,a、b、c、d 和 e 表示不同的肽链)(图 2-32)。如参与神经肌肉通信的乙酰胆碱受体是由 5 条肽链组成的膜蛋白(a_2bcd)。

由一条以上肽链组成的蛋白质称为寡聚体(oligomer)蛋白,其中的每一条肽链称为单体(monomer)或亚基(subunit)。含有 2、3、4、5 或 6 个亚基的蛋白质分别称为二聚体(dimer)、三聚体(trimer)、四聚体

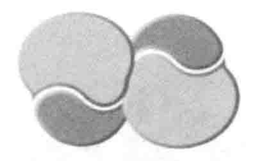
(tetramer)、五聚体(pentamer)或六聚体(hexamer)蛋白,含有更多亚基的蛋白质依此类推。最常见的寡聚体蛋白质是二聚体。由同一种肽链组装而成的寡聚体称为同源寡聚体,否则称为异源寡聚体。



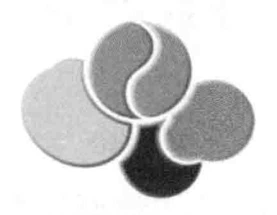
同源二聚体a_2



异源二聚体ab



异源四聚体a_2b_2



异源五聚体a_2bcd

图 2-32 构成蛋白质四级结构的亚基的不同组合(Petsko 等,2004)

构成异源寡聚体的各个单体尽管一般由不同的基因编码,且在某些情况下序列很少或没有同源性,但通常在三维结构上很像。这说明许多异源寡聚体起源于一个编码同源寡聚体的原始基因的重复。

(一) 四级结构的定义和内容

具有两条或两条以上多肽链的蛋白质如果不是以二硫键相连,则认为它们具有四级结构,其中的每一个亚基都有自己的三级结构。

蛋白质四级结构的内容包括亚基的种类、数目、空间排布以及亚基之间的相互作用。

由相同亚基组装的蛋白质通常是对称的。如果相互作用的亚基不一样,也就无对称可言。

相同的亚基为了相互作用,其中每一个亚基必须具有结合区域 A 和它的互补区域 A′,从而形成一种对称的复合物。受表面的 A 和 A′位置的决定,亚基相结合形成封闭的结构,以二聚体或三聚体最常见,或者形成非常少见的末端开放的链状结构(螺旋排列最常见)。组装成的对称复合物的重复单位可能是单体,也可能是由不同肽链组成的复合体。例如,血红蛋白由 4 条肽链组成($\alpha_2\beta_2$),为对称的二聚体,其重复单位为 αβ。在对称复合物中的对称单位称为原体(protomer)。如果亚基上含有第二套互补结合的区域(如 B 和 B′),那么它们还能形成更精妙的复合物,因为第二个结合区域使得对称结构能够继续配对。例如,配对的二聚体形成四聚体,配对的三聚体形成六聚体。

(二) 四级结构的形成

亚基的表面是不规则的,这使得亚基之间有可能彼此结合,并形成四级结构。而亚基之间是否真正结合不仅与其形状有关,而且与将亚基联系在一起的次级键有关,主要包括氢键供体对氢键受体,疏水基团对疏水基团,正电荷基团对负电荷基团。这些互补的性质在所有的结合作用中都可以观察到,无论是蛋白质和小分子之间,还是蛋白质和其他大分子之间。互补是必需的,因为分子间的界面由多个次级键组成,而互补性保证了两个结合的分子在界面上形成尽可能多的次级键。驱动四级结构形成或稳定四级结构的作用力包括氢键、疏水键、范德华力和离子键,这与一个单亚基蛋白稳定其内部折叠结构的键是一样的。虽然亚基之间的结合伴随着有序性的提高而导致熵的减少,这似乎在热力学上是不利的,但亚基结合使本来暴露在表面的疏水残基埋入到内部(亚基间的界面)却在热力学上又是有利的。研究表明,亚基之间的界面通常由约 60% 的疏水残基和 40% 的亲水残基组成。就一个典型的寡聚体蛋白而言,其两个亚基之间的解离常数 K_d 为 10^{-16}~10^{-8} mol/L,这说明了亚基之间的亲和力是相当高的。

非常稳定的寡聚体倾向于在亚基之间包埋更大的疏水表面,而聚合和解聚都很容易,并以此作为其功能一部分的亚基似乎利用更多的极性作用。亚基界面上存在很多氢键,大部分在亚基之间,其余在亚基和水分子之间。

不适当的四级结构会影响到某些蛋白质的功能,甚至会产生严重的功能障碍,像一些遗传疾病就起源于突变产生的不合适的疏水界面。例如,胰岛素是由 A 链和 B 链通过二硫键相连的一种激素。胰岛素只有以这种单体形式存在,才能与靶细胞膜上的受体结合,继而引发生物学效应。然而在溶液中,胰岛素能够自发地形成二聚体,二聚体之间还能聚合成六聚体。由于胰岛素分子形成六聚体的结合面就是受体与单个胰岛素分子结合的地方,因此,六聚体的胰岛素是没有活性的。从胰岛分泌出来的胰岛素是单体,它能够迅速作用于靶细胞。但是,如果将胰岛素通过静脉注射到糖尿病患者体内,原来形成的六聚体解离得较慢,致使降血糖效应来得慢,而影响到其药效。

1988 年,Guy Dodson 使用基因工程的方法对胰岛素进行了定点改造,将胰岛素缔合表面上的一个 Pro 残基用 Asp 取代。结果发现,界面上带负电荷的 Asp 通过静电斥力阻止亚基之间的结合,从而提高了六聚体解离成单体的解离常数。将这种胰岛素变体注射到实验动物体内比一般的胰岛素能更快

地诱发降血糖效应。

（三）四级结构形成的意义

具有四级结构的蛋白质具有特殊的优势，总结起来有如下 8 点：

（1）通过减少蛋白质表面积和体积的比率而提高蛋白质的稳定性。

（2）提高遗传学上的经济与效率　如果有大小相同的两种蛋白质，一种只有一条多肽链组成，另一种由两个相同的亚基组成，显然编码后者比前者所需要的碱基序列要少，而且表达的效率更高。

（3）有利于某些酶的活性中心的组装　如果蛋白质是酶，形成寡聚体可将催化位点聚集在一起。许多酶是从亚基的聚合获得催化或部分催化能力的。有时，单个亚基可能建立不了一个完整的活性中心，只有形成寡聚体才能将所有必需的催化基团集中在一起，组成一个有活性的酶。例如，大肠杆菌的谷氨酰胺合成酶（glutamine synthetase）的活性中心是由两个成对的亚基组装而成的，而解离的单个亚基没有活性。

（4）能产生协同效应（co-operativity）（参看第三章有关血红蛋白的结构与功能）。

（5）减少翻译出差错的机会　因为与同样大小的单体蛋白相比，构成寡聚体蛋白质的肽链更短，翻译出错的机会就更小（参看第三十九章“蛋白质的生物合成及其在细胞内的降解”）。

（6）改变一种蛋白质功能的特异性　例如，细菌的 RNA 聚合酶由 σ 因子和核心酶组成，只有结合有 σ 因子的核心酶才能识别启动子，启动基因的转录（参看第三十六章“DNA 转录”）。

（7）有利于酶活性的调控　许多酶由调节亚基和催化亚基组成（如蛋白质激酶 A），当调节亚基与特殊的配体分子结合以后，会影响到催化亚基的活性。还有些酶具有单体（无四级结构）和寡聚体（具有四级结构）两种形式，但只有一种形式有活性。

（8）有利于包装成更加庞大的结构　例如，病毒的衣壳通常由同一个相同的亚基自组装而成。

第三节　蛋白质的折叠历程与结构预测

既然蛋白质的高级结构直接决定其功能，那么如果能够深入了解蛋白质的折叠机制，将会大大促进对蛋白质的功能研究，在此基础上还可以对一些未知蛋白质的结构和功能进行合理的预测。如今科学家对蛋白质折叠的一些基本规律已有所了解，但对蛋白质折叠的详细机制的认识还很不完善。

一、蛋白质折叠的基本规律

到目前为止，已经被确认的蛋白质折叠的基本规律包括：

（1）一级结构决定高级结构，即一级结构包含了一种蛋白质折叠成最终构象所需要的全部信息，这是美国科学家 Anfinsen 早在 1954 年就已得出的结论（参看本章的科学故事）。

就某些蛋白质而言，这些折叠信息是自我识别的，因而它们不需要其他蛋白质的帮助就可以正确地折叠；而另外一些蛋白质，这些信息可能也存在，但需要其他蛋白质的帮助以防止错误的折叠；还有一些蛋白质，如枯草杆菌素（subtilisin），以前体的形式被合成，在后加工的过程中被切除的肽段充当肽链剩余部分折叠的模板。前枯草杆菌素原是枯草杆菌素的前体，它在 N 端含有两段前肽序列，一段为信号序列，指导其分泌到胞外，另一段作为折叠的模板。在后加工的时候，先切除信号序列产生枯草杆菌素原，然后再切除第二个肽段，产生有活性的枯草杆菌素。枯草杆菌素原的变性是可逆的，但枯草杆菌素则不然，这是由于枯草杆菌素已丧失了自我折叠的模板。

（2）蛋白质的折叠伴随着自由能的降低，$\Delta G_{折叠}(G_N-G_U)$ 为 −80~−20 kJ/mol（N 表示蛋白质最后的天然构象状态，U 表示非折叠状态），这属于热力学上有利的反应，但是两种状态在自由能上的差距并

不很大。小的自由能差异是必要的，因为如果差别太大，就意味着折叠后的蛋白质非常稳定，即刚性太强、柔性不足，这反而不利于蛋白质的功能发挥。

(3) 蛋白质的折叠是协同和有序的过程。

(4) 驱动蛋白质（特别是球状蛋白质）折叠的主要作用力是疏水键，其他次级键也有作用。就氢键而言，要尽可能让主链肽基之间形成最大数目的分子内氢键（如 α 螺旋、β 折叠），同时保持大部分能参与形成氢键的 R 基团处于蛋白质分子表面，与水相互作用。

(5) 在细胞内，不同的蛋白质可能具有不同的折叠路径（图 2-33）。

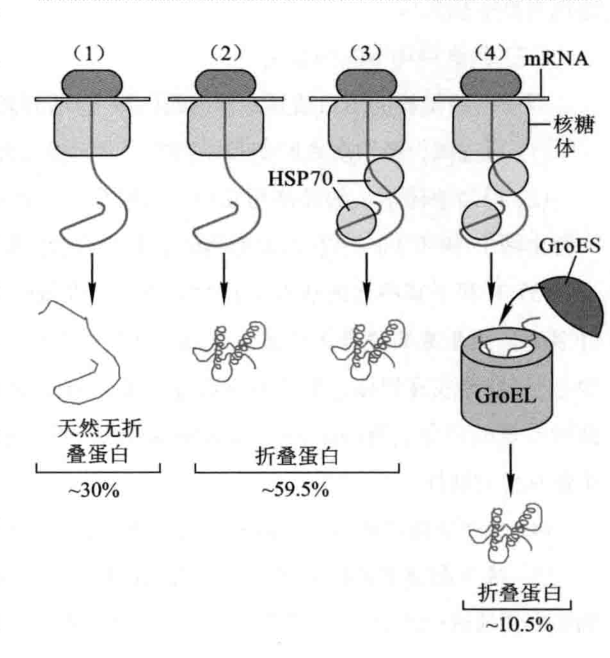


图 2-33　蛋白质折叠的不同途径

有些蛋白质在合成好以后，并不折叠或仅仅部分折叠，它们缺乏特定的二级结构和三级结构，处于完全无折叠或部分无折叠状态，但仍然具有功能（参看第三章有关天然无折叠蛋白质的结构与功能的内容）；有些蛋白质能够完全独立地进行折叠；有些需要其他蛋白质的帮助。蛋白质在寻找其天然的低能状态的时候，会不断地经历折叠和重折叠反应，途中需要经过多个高能的构象状态。此阶段构成蛋白质折叠的动力学障碍（kinetic barrier），而成为折叠的限速步骤。事实证明，细胞内有一类专门的统称为分子伴侣（chaperones）的蛋白质可用以克服这些动力学障碍，以加快折叠的速率，而克服动力学障碍的动力是 ATP 的水解。

分子伴侣在与待折叠的蛋白质结合以后，先稳定它们非天然的构象，然后再促进正确的折叠。这些待折叠的蛋白质可能刚刚从核糖体释放出来，或者正在跨膜但还没有折叠，或只是部分折叠。分子伴侣不会与已折叠好的处于天然构象状态下的蛋白质结合，也不会与细胞中天然的无结构蛋白质（natively unstructured protein，NUP）结合。多数分子伴侣是非特异性的，这意味着它们能够作用于一系列不同种类的肽链，而少数分子伴侣只能作用于特定的目标肽链。

分子伴侣所起的作用主要是在正确的时间和正确的地点促进新生肽链的正确折叠，同时防止它们发生降解和沉淀，有时还能“拨乱反正”，帮助错误折叠的蛋白质有机会重新折叠成正确的构象。分子伴侣并非“终身伴侣”，一旦蛋白质折叠好，它就被释放，然后再参与另一个新生肽链的折叠。常见的分子伴侣有两大类：热激蛋白 70 类家族（heat-shock protein 70，HSP70）和伴侣蛋白（chaperonin）家族。

HSP70 类分子伴侣在细胞内与新生肽链上的疏水区域临时结合，一方面能阻止多肽链提前折叠；另一方面可在多肽链没有折叠之前，防止暴露在外的疏水区之间通过疏水作用“非法聚集”在一起，而导致多肽链之间的聚合，甚至出现沉淀而伤害细胞。伴侣蛋白则形成大的桶状结构，来容纳部分折叠的蛋白质完成折叠［图 2-33(4)］。

根据结构和作用机制，伴侣蛋白还可分为两类：细菌的伴侣蛋白属于第一类，古菌的伴侣蛋白和真核生物的 TRiC 伴侣蛋白属于第二类。以大肠杆菌的 GroEL-GroES 复合物为例，它由 GroEL 和 GroES 组成，其中 GroEL 单个亚基的 M_r 为 6×10^4，共有 14 个亚基。这些亚基组装成两层七元环桶状结构，GroES 充当桶顶部的圆盖［图 2-33(4)］，可以和 GroEL 分开。新生肽链进入圆筒内部以后，顶盖合上，肽链在内部与桶壁发生多次可逆结合，每一次结合由 ATP 水解所驱动，在最终折叠成特定的三维结构之前可能会消耗大量的 ATP。一旦折叠好，便从 GroES-GroEL 复合物中释放出来。比如，一个硫氰酸酶（rhodanese）分子，在圆筒内达到完全折叠状态需要消耗 130 个 ATP。第二类伴侣蛋白的顶盖是内置的，由顶部亚基突出的结构域组装而成。

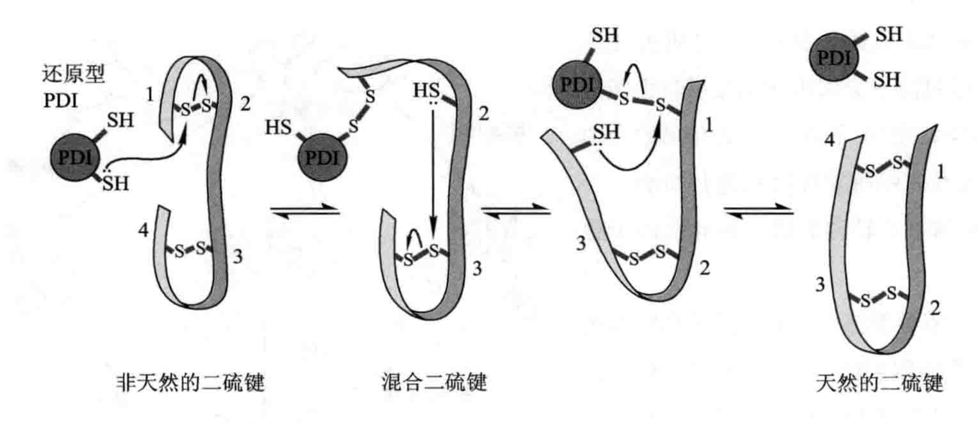


图 2-34　PDI 的作用机制（Voet 等，2011）

（6）某些蛋白质的折叠还需要蛋白质二硫化物异构酶（protein disulfide isomerase，PDI）和肽酰脯氨酰顺反异构酶（peptidylprolyl *cis-trans* isomerase，PPI）的帮助。

PDI 的功能是促进含有二硫键的蛋白质形成正确的二硫键。其作用的机理是通过重排二硫键，使蛋白质能快速找到热力学最稳定的二硫键配对方式。

PDI 一般含有两段 CGHC 序列，处于这种序列中的 Cys 的巯基具有较低的 pK_a(7.3)，低于其他蛋白质分子中巯基的 pK_a(8.5)，故它的反应性较强，在生理 pH 下，就能够充当良好的亲核试剂，从而启动二硫键的重排反应。PDI 的作用机制如下：首先，它的一个反应性强的巯基进攻暴露在外的错误的二硫键，形成混合二硫键；随后发生二硫键的重组，直至形成正确的二硫键。由于正确的二硫键处于正确的三维结构之中，PDI 很难再对其进攻，所以被保留下来（图 2-34）。

真核细胞内的 PDI 位于内质网，那些分泌到胞外的或者最后定位到质膜的蛋白质在分拣和定向的过程中（参看第四十章“蛋白质的翻译后加工及其定向和分拣”）途经内质网，因此能够形成二硫键。而那些在细胞质游离的核糖体或者在细胞器（叶绿体和线粒体）核糖体上完成翻译的蛋白质（如核蛋白、细胞质基质蛋白、线粒体蛋白和叶绿体蛋白）是一般不会含有二硫键的。细菌的 PDI 又称为二硫桥形成蛋白（disulfide bridge-forming protein，Dsb），它位于细胞外的周质（periplasm），因此只有分泌到周质中的蛋白质才有可能形成二硫键。

PPI 的功能是促进 X—Pro 之间的肽键采取正确的形式。蛋白质分子中在 X—Pro 之间的肽键大概有 6% 以顺式构型存在。X—Pro 之间肽键的顺反异构构成许多蛋白质折叠的限速步骤。PPI 通过扭曲这个位置的肽键来促进它的顺反异构，从而加速蛋白质的折叠。Pin1 蛋白就是人体内一种重要的 PPI，它在细胞内能够特异性地识别磷酸化的 S/T-P 序列，催化其中的肽键发生顺反异构，从而调节细胞内一系列蛋白质的活性，特别是一些与细胞周期调控有关的蛋白质（如周期蛋白 D1）。已发现，乳癌细胞内的 Pin1 呈过量表达。

（7）最终得到的蛋白质构象不是僵硬的，而是具有一定的柔性。根据 X 射线衍射或 NMR 绘制出来的蛋白质构象图是僵硬和静止的，但实际上蛋白质是高度柔性的分子，也许用“刚柔相济”这个成语描述蛋白质的结构特征更形象。

e2-2 浅谈蛋白质结构的柔性

二、蛋白质折叠的历程

在 Anfinsen 提出蛋白质折叠的热力学假说以后，人们对于蛋白质在折叠过程中最后如何达到在热力学上最稳定的构象状态有很大的争议。1968 年，Cyrus Levinthal 对此提出了所谓的 Levinthal 氏悖论（Levinthal's paradox）：假定一种蛋白质由 100 个氨基酸残基组成，而每个残基有 2 个构象，那么这种蛋白质的总构象数目是 $2^{100}=1.27\times10^{30}$。再假定此蛋白质为了寻找总能量最低的构象状态，每尝试一次构象耗时 10^{-13} s，那么将所有可能的构象都尝试完需要的时间是 $(10^{-13})(1.27\times10^{30})=1.27\times10^{17}$

秒 $=4\times10^9$ 年！

Levinthal 氏悖论显示，蛋白质折叠不可能是通过随机尝试来找到最后自由能最低的构象状态的，而是只尝试有限数目的构象，或者按照特异性途径进行折叠。于是，研究的重心转移到揭示特异性的折叠途径上。

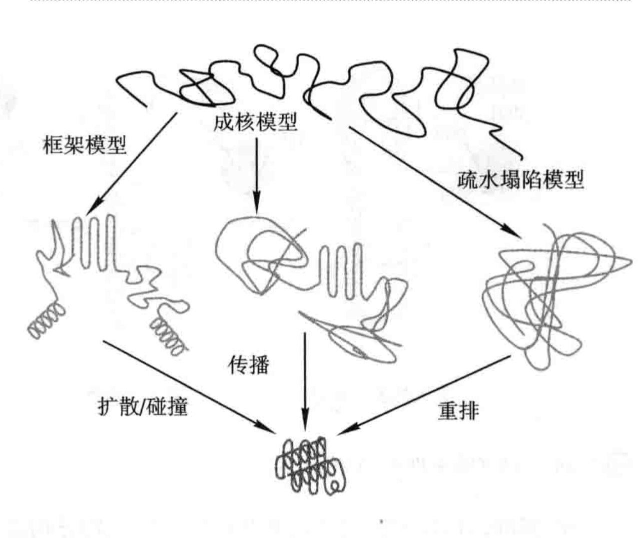


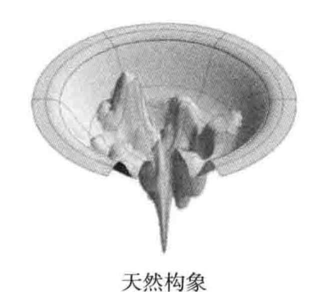
图 2-35　蛋白质折叠的三大模型

科学家根据体外实验的结果和计算机模拟等手段得到的数据，提出了 3 种模型用来解释蛋白质的折叠过程（图 2-35）：

（1）框架模型（framework model）　该模型认为，局部的二级结构首先形成，它们独立于三级结构的建立。当折叠好的各种二级结构单元扩散并发生碰撞的时候，便发生了聚合，从而成功地形成三级结构。

（2）疏水塌陷模型（hydrophobic-collapse model）　该模型认为，蛋白质分子上的疏水侧链快速地发生包埋，即发生疏水塌陷，亲水侧链则暴露在外，形成熔球体（molten globule）。在熔球体内，远距离基团之间的相互作用得以建立，从而先形成三级结构，最后才形成二级结构成核模型。

（3）成核模型（nucleation model）　该模型认为，在一级结构上相邻的一些序列自发折叠成天然的二级结构，然后这些二级结构充当折叠核（folding nuclei），其他结构以此为核心，向周围扩展，逐步形成最终的三级结构。



有人将这 3 种模型结合起来，认为蛋白质折叠经过 3 步反应：①启动，快速地形成局部二级结构，即折叠核，此过程是可逆的；②折叠核协同聚合成结构域；③结构域经熔球体，最终形成具有完整三维结构的蛋白质。熔球体被认为是疏水塌陷的结果，这种中间体含有某些二级结构，但还没有形成正确的三级结构。

如果用图 2-36 来表示上述折叠过程，可能更为形象。蛋白质折叠的全景图好像一个漏斗，漏斗表面的每一个点代表的是多肽链的一个构象，蛋白质可被视为一组在向下滑行的滑雪者，下坡滑雪指导多肽链进入它的天然状态。有些蛋白质折叠过程比较"坎坷"，有些蛋白质则折叠得比较顺利。蛋白质的天然状态是最深的"山谷"，其他山谷代表的是部分折叠的构象状态。

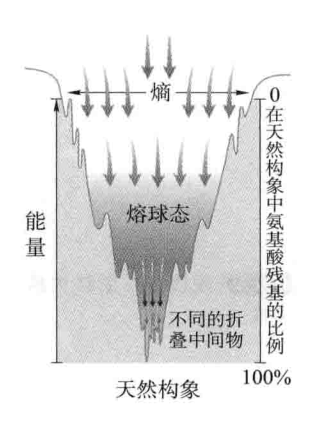


图 2-36　蛋白质折叠的"滑雪"模型

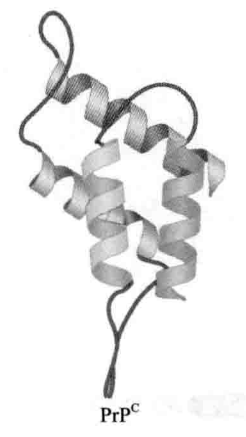
一般认为，蛋白质在溶液中折叠的驱动力包括：①肽链内氢键供体和受体之间形成氢键，而将原来结合的水分子释放，从而增加水的熵；②疏水侧链倾向聚合以尽可能减少与水接触的疏水面积；③亲水残基面向表面与水接触，增加溶解性；④带相反电荷的基团靠近，形成盐桥；⑤在疏水作用力的基础上，疏水基团靠近而产生范德华引力。这些作用力也是稳定蛋白质三维结构的因素，而不利于蛋白质三维结构稳定的因素则包括：①肽链形成有规律的二级结构，使熵减少；②将一个极性的基团从水相中移走，却没有让它形成新的氢键；③将一个带电基团从水相中移走，却没有带相反电荷的基团和它形成盐键，或者将两个带同种电荷的基团强行放在了一起；④将一个疏水基团放在水相中；⑤将两个原子放在同一位置，产生空间位阻。也许，蛋白质折叠就是这两种相反的势力相互较劲最后妥协的结果，仿佛是一个竞争激烈又相互谦让的游戏。例如，牺牲 1~2 个氢键却换来了 2~3 个疏水键，或者将 1 个疏水基团与水接触却让其二级结构多了一些氢键。

三、与蛋白质错误折叠相关的疾病

据估计，正常的细胞约有三分之一的蛋白质可能会发生错误折叠，但细胞内有专门的质量控制系统能及时发现并处理它们。例如，分子伴侣能够与错误折叠的中间物结合并重启折叠过程。另外，一

种称为泛素(ubiquitin)的蛋白质能够将这些错误折叠的蛋白质打上"死亡"标签,并把它们带到细胞内的"垃圾处理站"——一种被称为蛋白酶体(proteasome)的圆筒状细胞器进行水解,以防止它们在细胞内的堆积(参看第三十九章"蛋白质的生物合成及其在细胞内的降解")。因此,一般情况下细胞内出现少量折叠异常的蛋白质并不会影响到细胞的正常功能。然而,如果一个细胞内大量出现某种错误折叠的蛋白质,以至于超过了蛋白酶体处理能力的时候,就可能导致机体的病变。蛋白质错误折叠最麻烦的问题就是它们可能会聚集在一起,最终形成不溶性的淀粉样纤维(amyloid fibril)损害细胞,严重的可导致细胞死亡。

近些年来,人们发现越来越多的疾病与蛋白质的异常折叠有关,例如囊性纤维变性(cystic fibrosis, CF)和一些神经退化性疾病,其中囊性纤维变性是由于基因突变导致编码的蛋白质不能正确折叠,而神经退化性疾病主要是因为后天因素导致神经细胞内出现一些细小的原纤维(protofibril)。原纤维由4~30个错误折叠的蛋白质形成,它们可进一步聚合成不溶性的纤维,而不论是原纤维,还是纤维,都会危害细胞。



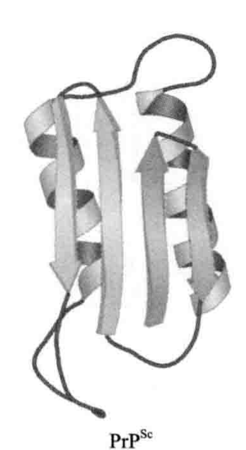


图 2-37 PrP^C 与 PrP^{Sc} 在构象上的主要差别(Campbell 等,2009)

(一) 海绵状脑病(spongiform encephalopathy, SE)

SE 是一种致命性神经退化性疾病,因受感染的动物在脑部病变的部位出现海绵状的空洞而得名,它是由错误折叠蛋白引发的一代表性疾病。SE 可以感染多种动物,例如人的克雅氏病(Creutzfeldt-Jakob disease, CJD)、GSS 氏病(Gerstmann-Straussler-Scheinker disease)、库鲁病(Kurudisease)、致死性家族性失眠症(fatal familial insomnia, FFI)、幼儿海绵状脑病(alpers disease)、山羊和绵羊的羊瘙痒症(scrapie)、鹿和麋的慢性消耗病(chronic wasting disease)以及牛的海绵状脑病,即疯牛病(mad cow disease)。SE 的主要症状包括渐进性痴呆(progressive dementia)和运动机能失调。

SE 的致病因子是一种折叠异常的朊蛋白(prion protein scrapie, PrP),通常简写为 PrP^{Sc}。正常动物也有这种蛋白质,一般简写成 PrP^C(prion protein cellular, PrP^C)。尽管 PrP^C 与 PrP^{Sc} 的一级结构完全一样,但构象不同。PrP^C 与 PrP^{Sc} 的主要差别是(图 2-37):PrP^C 富含 α 螺旋,可溶于水,对蛋白酶敏感,呈单体状态;PrP^{Sc} 则富含 β 折叠,其核心部分能抵抗蛋白酶的水解,分子间很容易聚合形成多亚基聚合体,最后形成淀粉样纤维杆状结构。

PrP^C 主要分布在脑细胞上,少量发现在淋巴细胞和其他细胞中。PrP^C 的确切功能尚不清楚,但它在进化上十分保守。成熟的 PrP^C 由 209 个氨基酸残基组成,含有一个二硫键,通过糖基化的磷酸肌醇(glycosylphosphatidylinositol, GPI)锚定在细胞膜的外侧。人的 PrP^C 由位于 20 号染色体短臂上的 *PRNP* 基因编码。

动物患病途径有 3 条,即受外来的朊病毒感染、家族性遗传和 PrP^C 偶然的折叠错误。无论是哪一种方式,都是先出现少量的"坏的"蛋白质——PrP^{Sc},而 PrP^{Sc} 一旦出现,自身可以作为模板,催化细胞内原来"好的"蛋白质——PrP^C 向"坏的"PrP^{Sc} 转变,由此导致神经退化和病变,真可谓是"近墨者黑"。烹调和煮沸都不能破坏朊病毒,因此,感染朊病毒的动物是绝对不能加工食用的,必须将其彻底销毁。

Quiz12 根据 PrP^{Sc} 致病的机制,试设计一种治疗 CJD 的方法。

通过同源重组技术可以获得 *PRNP* 基因被敲除的小鼠,这种小鼠似乎能正常地发育和生殖,但却失去了感染朊病毒的能力。通过进一步追踪研究发现,缺失 *PRNP* 基因的小鼠还是表现出了一些异常,如外周神经脱髓鞘、对缺氧性脑损伤敏感性增强、睡眠和昼夜节律受到影响、年老的时候可得小脑共济失调等。

PrP^C 与 PrP^{Sc} 被认为具有相同的能量状态。幸运的是,PrP^C 自发重折叠成 PrP^{Sc} 的可能性很低,这是因为两者的转变需要非常大的活化能。因此,在正常人的一生中,PrP^C 几乎不可能形成 PrP^{Sc}。

家族型朊病毒疾病由 PrP^C 基因突变造成。突变降低了 PrP^C 重折叠成 PrP^{Sc} 的活化能,使 PrP^C 形成 PrP^{Sc} 的机会大增。据估计,正常细胞因折叠错误产生 PrP^{Sc} 需要三四千年,而突变引起的活化能降低则使时间缩短到 30~40 年,正好落在一个人正常的生命周期内。

Quiz13 你认为 PrP^C 基因的突变是显性的还是隐性的?

(二) 囊性纤维变性

囊性纤维变性是一种遗传性疾病,它与一种编码囊性纤维变性跨膜传导调节蛋白(cystic fibrosis transmembrane conductance regulator,CFTR)的基因缺陷有关。但致病的原因并不是患者体内有缺陷基因编码的异常蛋白,而是缺乏正常的 CFTR 才致病的,因此这种疾病是一种隐性遗传病,即编码 CFTR 的两个等位基因必须同时有缺陷才能致病。

正常的 CFTR 是一种跨膜蛋白,存在于上皮细胞的质膜上,作为一种受 ATP 调控的氯离子通道,负责将氯离子从细胞内泵出去。有缺陷的 CFTR 不能正常地折叠,其折叠中间物不能与分子伴侣解离,从而导致它不能整合到细胞膜上。于是,缺乏 CFTR 的上皮细胞开始积累氯离子。由于渗透压的缘故,细胞开始拼命吸收周围黏液中的水分,导致黏液变稠,特别在患者的呼吸道内会产生稠厚的黏痰,而阻碍呼吸,引起阵阵剧烈咳嗽,使得呼吸道反复受感染。患者平均期望寿命仅为 37 岁。

除了海绵状脑病和囊性纤维变性以外,阿尔茨海默氏病(Alzheimer's disease,AD)和帕金森病(Parkinson's disease,PD)也是两种因蛋白质错误折叠引起的疾病。

e2-3 AD 和 PD

四、蛋白质结构的预测

随着基因组学的发展,越来越多的物种的基因组序列被测定出来。然而面对浩瀚的基因组数据,科学家最大的挑战是,它们决定的蛋白质功能是什么?既然结构决定功能,要想确定一个蛋白质的功能,首先就需要确定它的结构,特别是三维结构。确定一个蛋白质的一级结构也许并不困难,通过其基因的核苷酸序列就能很容易推断出来,但要确定其三维结构就不那么容易了!

既然一级结构决定蛋白质高级结构,那么能否通过一级结构来预测它的高级结构呢?可以说,蛋白质结构的预测被视为生物信息学的"圣杯"(Holy Grail),更是科学家的梦想。

(一) 二级结构的预测

与预测蛋白质的三级结构相比,二级结构的预测要容易一些。首先是因为二级结构是一种局部结构,其次是不同的氨基酸对不同的二级结构具有明显的倾向性。

如表 2-2 所示,在 20 种常见的蛋白质氨基酸残基中,有的适合形成 α 螺旋,有的正好相反,甚至

表 2-2 不同的氨基酸残基形成 α 螺旋、β 折叠和 β 转角的趋势

氨基酸	α 螺旋(P_α)	β 折叠(P_β)	β 转角(P_t)	有利于形成 α 螺旋	有利于形成 β 折叠	有利于形成 β 转角
Ala	1.29	0.90	0.78	√		
Cys	1.11	0.74	0.80	√		
Leu	1.30	1.02	0.59	√		
Met	1.47	0.97	0.39	√		
Glu	1.44	0.75	1.00	√		
Gln	1.27	0.80	0.97	√		
His	1.22	1.08	0.69	√		
Lys	1.23	0.77	0.96	√		
Val	0.91	1.49	0.47		√	
Ile	0.97	1.45	0.51		√	
Phe	1.07	1.32	0.58		√	
Tyr	0.72	1.25	1.05		√	
Trp	0.99	1.14	0.75		√	
Thr	0.82	1.21	1.03			√
Gly	0.56	0.92	1.64			√
Ser	0.82	0.95	1.33			√
Asp	1.04	0.72	1.41			√
Asn	0.90	0.76	1.28			√
Pro	0.52	0.64	1.91			√
Arg	0.96	0.99	0.88			

是 α 螺旋的“终结者”；有的适合形成 β 折叠或 β 转角。例如，Gly—Pro 和 Pro—Gly 序列一般意味着转角的出现。

综合各种氨基酸残基形成不同二级结构的趋势，P. Y. Chou 和 G. D. Fasman 在 1974 年曾提出预测球状蛋白质二级结构的一套规则：①由≥6 个氨基酸残基组成的肽段，如果其中的 $P_{\alpha} \geq 1.03$，且 $P_{\alpha} > P_{\beta}$，此外还没有 Pro，那么这个肽段应该形成 α 螺旋；②由≥5 个氨基酸残基组成的肽段，如果 $P_{\beta} \geq 1.05$，且 $P_{\beta} > P_{\alpha}$，那么这个肽段应该形成 β 折叠；③如果一个四肽序列的 $P_{\alpha} < 0.9$，且 $P_{t} > P_{\beta}$，那么它很可能形成 β 转角。

对于两种主要的二级结构而言，α 螺旋也许可被视为多肽链的默认二级结构状态，尽管其势能不比 β 折叠低，但其氢键形成在螺旋内。而 β 折叠的氢键在股与股之间形成，参与形成氢键的 β 股可能在一级结构上相距较远，因此 α 螺旋与 β 折叠相比，具有熵的优势。成簇的 α 螺旋或 β 折叠的“破坏者”会形成环或转角结构，它们作为有规则二级结构的边界，将相邻的二级结构联系起来。

尽管二级结构的预测不可能绝对可靠，但总的趋势是存在的。目前，结合几种统计学途径得出的一系列规则使得二级结构预测的准确率一般能达到 80%。例如，有人从人类基因组中得到的一个未知功能的蛋白质的一段氨基酸序列，单凭预测得到的二级结构就已取得很高的准确度。如图 2-38 所示，有 8 个统计学预测方法被用于此段序列。最引人瞩目的是，所有方法在 α 螺旋（用 h 表示）和 β 股（用 e 表示）出现的大概位置上是一致的，但在螺旋和股的长度以及末端的位置上有较大的分歧。此外，对环出现的可能位置（用 c 表示）也有争议。这样的结果不管怎样，还是颇为典型的。显然在进行预测的时候，使用多种方法要比只用一种方法更具有启发性。图 2-38 中的 Sec. Cons. 线显示的是一种序列预测，其下方是根据 X 射线晶体衍射得出的结果，α 螺旋和 β 股的实际位置分别用圆柱体和单向粗箭头表示，回折细箭头表示 β 转角。

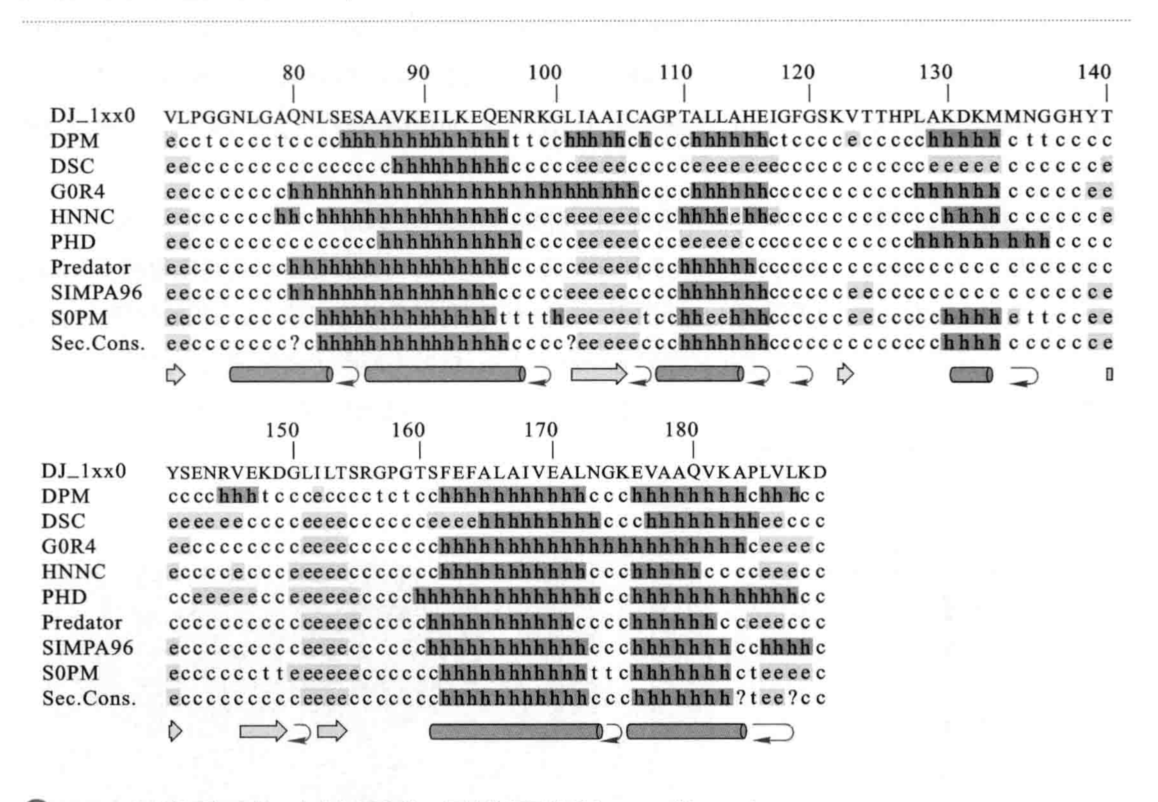


图 2-38　人类基因组中的一个未知蛋白的二级结构预测结果（Petsko 等，2004）

（二）三级结构的预测

至少到目前为止，使用从头计算法（ab initio）即单凭氨基酸序列来预测三维结构还很不成功。要想获得一种蛋白质精确的三级结构，仍然需要借助于 X 射线晶体衍射和 NMR。但由于这两种技术都各有自己的局限，因此永远也跟不上基因组测序的步伐。

为什么三维结构预测，尤其是使用从头计算法这么困难？这主要是因为蛋白质在折叠的时候，各个氨基酸残基有不同程度的自由度，并且驱动和稳定折叠的化学键很多是在一级结构相隔甚远的R基团之间的相互作用，此外，自然折叠并不经历所有的构象状态，也会有一些酶和分子伴侣协助。尽管现在有计算机协助，但一种蛋白质可能的构象状态也超过了当今最快的计算机的运算能力。也许蛋白质在折叠的过程中，只有几百个基本的结构，但现在我们还没有能力去识别它们。

第四节　蛋白质组及蛋白质组学

蛋白质组(proteome)是指所有基因表达的全部蛋白质及其存在方式，即一个基因组、一个细胞或组织所表达的全部蛋白质成分。

蛋白质组学(proteomics)则专门研究不同时间和空间发挥功能的特定蛋白质组，旨在阐明生物体全部蛋白质的表达模式及功能模式，其内容包括确定蛋白质表达、存在方式(修饰形式)、结构、功能和相互作用方式等，从一个机体或一个细胞的蛋白质整体活动来揭示生命规律。研究对象可以是细胞或组织的全部蛋白质，也可以是一个特定的生物学机制或与机制相关的全部蛋白质。总体上研究可以分为3个方面：①研究蛋白质表达模式即蛋白质组成。这需要对蛋白质进行大规模地分离与鉴定，并对其表达后修饰进行鉴定；②研究蛋白质功能模式，目前主要集中在蛋白质相互作用的网络关系。然而，由于蛋白质具有多态性、可变性、复杂性，并且低表达蛋白质难以检测等，进行相关的研究并非易事；③蛋白质的差异显示及其在疾病发生中的作用。

与基因组相比，蛋白质组具有多样性和可变性。对一个机体而言，基因的数目是恒定的，而蛋白质的种类和数目在同一个机体的不同类型的细胞中各不相同，即使是同一细胞，在不同的时期或不同条件下，其蛋白质表达也不同。由于蛋白质种类并不等于基因组中可读框的数目，因此有人提出在特定时间、特定环境和实验条件下基因组活跃表达的蛋白质为功能蛋白质组。显然，它们只是总蛋白质组的一部分。

目前，蛋白质组学研究以双向电泳和质谱(MS)分析两大技术为核心。其中双向电泳(图2-39)是以等电聚焦为第一向凝胶，根据蛋白质p*I*不同而将其分离。第二向凝胶是十二烷基硫酸钠聚丙烯酰胺(SDS-PAGE)凝胶电泳，根据蛋白质大小的不同来分离蛋白质(参看第四章“蛋白质的性质、分类及研究方法”)。质谱技术的基本原理是将样品分子离子化后，根据不同离子间质荷比的差异来分离并确定大小。除了质谱以外，还有蛋白质芯片、蛋白质序列分析和氨基酸组成分析等技术可用，但以质谱分析和蛋白质序列分析最为重要。

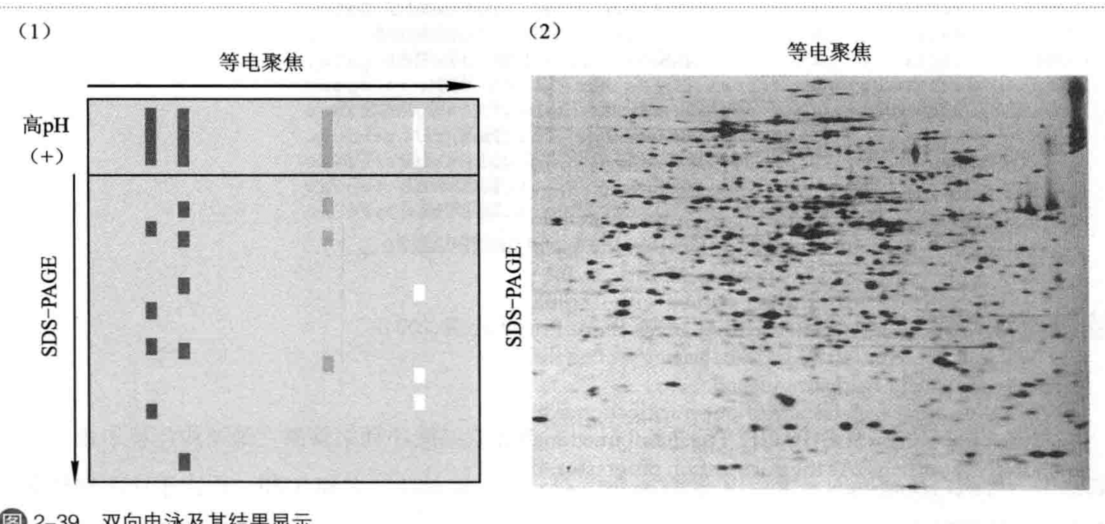


图2-39　双向电泳及其结果显示

科学故事——"一级结构决定高级结构"学说的提出

蛋白质的高级结构由其一级结构决定的学说最初由 Christian B. Anfinsen 于 1954 年提出。在 1950 年之前，Anfinsen 一直从事蛋白质结构方面的研究。在进入美国国立卫生研究所(NIH)以后，他继续从事这方面的研究。当时他最想知道的是：一个蛋白质分子是如何折叠成其独特的三维构象的？需要不需要其他酶的帮助？蛋白质为什么要采取特定的构象？

然而，要想了解蛋白质的折叠过程，首先需要建立一种方法能够测定蛋白质的构象，其次还需要找到一种手段用以检测折叠过程。Anfinsen 以牛胰核糖核酸酶(ribonuclease)为研究对象轻而易举地解决了第一个问题(实际上选用此酶的部分原因是当时的芝加哥肉类加工厂能够为他的实验室随时提供足够的原材料)，由于核糖核酸酶催化 RNA 的水解，其酶活性完全取决于其特定的三维构象，于是酶活性成为测量这种蛋白质采取何种构象的一种方法。但要观测折叠过程，既可以从一个新合成的尚没有折叠的蛋白质开始，也可以在体外将一个已折叠好的蛋白质去折叠(unfold)，然后再观察它的再折叠(refold)过程。Anfinsen 选择了后一种途径。事实上，这种酶特别适合用后一种途径进行研究，首先因为二硫键是决定其分子形状的主要因素。其次，此酶是一种单纯蛋白质，只有 124 个氨基酸残基组成，含有 4 个二硫键，其活性很容易通过测定水解 RNA 释放出来的核苷酸量来获得。Anfinsen 和两个博士后 Michael Sela 和 Fred White 在研究中发现(图 2-40)，使用高浓度的巯基试剂——β- 巯基乙醇(β-mercaptoethanol)可将二硫键还原成自由的巯基，如果再加入尿素，进一步破坏已被还原的核糖核酸酶分子内部的次级键，则该酶由去折叠状态转变成无任何活性的无规则卷曲。对被还原的核糖核酸酶的物理性质进行分析的结果清楚地表明，它的确采取的是无规则卷曲的形状。

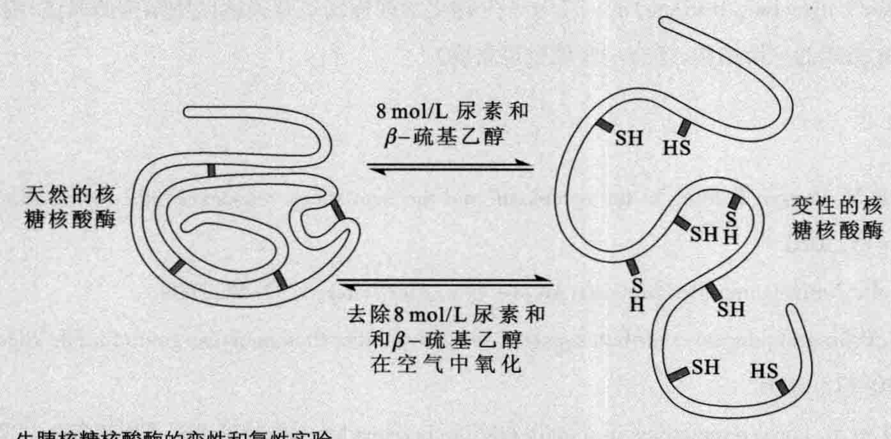


图 2-40　牛胰核糖核酸酶的变性和复性实验

在成功得到一种去折叠的核糖核酸酶以后，Anfinsen 便开始研究它的重折叠过程了。考虑到被还原的核糖核酸酶要在已被还原的 8 个 Cys 残基上重建 4 对二硫键，共有 105 种不同的组合，但只有一种是正确的形式，如果决定蛋白质构象的信息一直存在于氨基酸序列之中，那么最后重折叠得到的总是那种正确的形式。否则，重折叠将是随机的，最后只能得到少量的正确形式。显然，第一种情形能完全恢复去折叠过程中丧生的酶活性，而后一种情况只能恢复很少的酶活性。

Anfinsen 的重折叠实验还是比较顺利的，他通过透析的方法除去了导致酶去折叠的尿素和 β- 巯基乙醇，再将没有活性的酶转移到其生理缓冲溶液之中，在有氧气的情况下于室温放置，以使巯基能被重新氧化成二硫键。经过一段时间以后，发现绝大多数核糖核酸酶的活性得以恢复，这意味着它原来的构象恢复了。由于上述过程没有细胞内任何其他成分的参与，完全是一种自发的过程，因此有理由相信，此蛋白质正确折叠所需要的所有信息全部存在于它的一级结构之中。很快，Anfinsen 的研究成果被发表在 1954 年的 *Journal of Biological Chemistry* 上。进一步的实验确定了变性的核糖核酸酶完全恢复其活性的条件，在此基础上，Anfinsen 提出了蛋白质折叠的热力学学说(thermodynamic hypothesis)。此学说认为，一个蛋白

质的天然三维构象对应于它在生理条件下所处的热力学最稳定的状态。热力学稳定性是由组成蛋白质的氨基酸残基之间的相互作用决定，于是蛋白质的三维构象直接由它的一级结构决定。

尽管 Anfinsen 的工作奠定了蛋白质折叠的热力学基础，他也因此得到了 1972 年的诺贝尔化学奖，但是蛋白质折叠是相当复杂的，到现在为止，我们仍然不能根据一个蛋白质的一级结构推断出它的三维结构。同时，还必须注意到，蛋白质在体外的折叠比在细胞内的折叠要慢得多！

e2-4 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Protein（维基百科有关蛋白质部分的内容）
2. http://themedicalbiochemistrypage.org/protein-structure.html（完全免费的医学生物化学课程网站有关蛋白质结构的内容）
3. http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/P/Prions.html（内有各种朊病毒相关的内容，信息丰富）
4. http://www.friedli.com/herbs/phytochem/proteins.html（虽然是一个与草药化学有关的网站，但对蛋白质结构的描述非常翔尽和到位）
5. http://www.web-books.com/MoBio/Free/Chap2.htm（完全免费的分子生物学课程网站，有关蛋白质结构与功能的内容简单明了）
6. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb1/part2/protein.htm（美国伦斯勒理工学院——（Rensselaer Polytechnic Institute）的一个分子生物化学课程网站有关蛋白质结构的概述，内有多种蛋白质 PDB 格式的三维结构，还有一些题目可自测）

参考文献：

1. Vabulas R M. Protein Folding in the cytoplasm and the heat shock response. *Cold Spring Harb Perspect Biol*, 2: 1–15, 2010.
2. Nishikawalk. Natively unfolded proteins: An overview. *Biophysics*, 5: 53–58, 2009.
3. Pace C N. Polar group burial contributes more to protein stability than nonpolar group burial. *Biochemistry*, 40(2): 310–313, 2001.
4. Rose G D, et al. Local interactions in protein folding: Lessons from the α-helix. *J Biol Chem*, 272: 1 413–1 416, 1997.
5. Anfinsen C B. Principles that govern the folding of protein chains. *Science*, 181: 223–230, 1973.

第三章　蛋白质的功能及其与结构之间的关系

蛋白质是生物体各项功能的执行者。然而,任何一种蛋白质的功能都与其特定的结构密不可分,特别是三维结构。研究蛋白质结构与功能的关系是当今蛋白质组学最重要的内容之一,而根据一级结构的信息预测一种多肽或蛋白质的高级结构,并进而对其功能进行预测一直是科学家们的终极目标。

本章将重点介绍蛋白质的功能及其与结构之间的关系,以及几种重要的蛋白质的结构与功能。

第一节　蛋白质的功能

在某种意义上,每一种蛋白质都可视为一种独特的生物功能试剂,因为生物体内的每一项功能几乎都涉及到一种或几种甚至许多种特定的蛋白质。对一种蛋白质功能的生物学定义可以从不同的角度来理解。

对于生物化学家来说,一种蛋白质的功能意味着其在机体内承担的生化角色。如果是酶,其功能就是催化反应;如果是信号分子或运输蛋白,其功能就是在信号转导或运输途径中与其他分子之间发生相互作用。对于遗传学家或细胞生物学家来说,蛋白质的功能不仅包括其生化功能,还包括它的细胞学功能,例如它作用的途径,它的缺失或突变引起的表型变化。对于生理学家或发育生物学家来说,蛋白质功能的定义则更为广泛。

这里只集中讨论蛋白质的生化功能,它们主要包括:

(1) 充当生物催化剂　即酶,催化细胞内的各种生化反应(参看第八章“酶学概论”)。

(2) 调节其他蛋白质行使特定的生理功能或者调节基因的表达　例如,周期蛋白(cyclin)调节依赖于周期蛋白的蛋白质激酶(cyclin-dependent protein kinase,CDK)的活性;阻遏蛋白(repressor)和激活蛋白(activator)分别抑制和激活特定基因的表达。

(3) 运输　例如,血红蛋白运输氧气,转铁蛋白(transferrin)运输 Fe^{3+},白蛋白运输脂肪酸,载脂蛋白运输脂肪和胆固醇,玉米脂质转移蛋白(lipid-transfer protein)运输油酸。

(4) 贮存　例如,铁蛋白(ferritin)为细胞贮存铁,肌红蛋白为肌肉细胞贮存氧气。

(5) 运动　例如,肌动蛋白(actin)和肌球蛋白(myosin)的相互滑动导致肌肉细胞收缩或松弛。

(6) 为细胞和机体提供结构支持　例如,胶原蛋白在动物的结缔组织中就起结构支持的作用。

(7) 信号转导　例如,胰岛素及其受体的相互作用导致血糖浓度的下降(参看第十七章“激素与受体介导的信号转导”)。

(8) 免疫　例如,抗体参与体液免疫,T 细胞受体参与细胞免疫。

(9) 产生特定的毒性　例如,蓖麻毒素抑制真核细胞的蛋白质合成。

(10) 具有一些奇异的功能　例如,来自维多利亚水母体内的绿色荧光蛋白受紫外线的激发,可发出绿色的荧光;一种植物的乐果甜蛋白(monensin)具有极高的甜度,是一种天然的甜味剂;来自南极鱼体内的抗冷冻蛋白可帮助南极鱼抵御严寒;来自某些节肢动物的弹性蛋白(resilin)具有超常的弹性;由某些海洋动物(如贻贝)分泌的胶蛋白具有超强的黏性。

然而,并不是一种蛋白质只能行使一种功能,近来发现一些蛋白质虽然只有一种结构,但却能行使多种不同的功能。这些兼有其他功能的蛋白质被称为兼职蛋白(moonlighting protein)。例如,磷

酸己糖异构酶除了在细胞内参与糖酵解以外，还能由T淋巴细胞分泌到胞外充当一种神经白介素（neuroleukin），促进胚胎内某些脊髓神经元和感觉神经元的存活，以及促进B淋巴细胞的成熟。此外，它还是一种自分泌运动因子（autocrine motility factor），在由某些癌细胞分泌以后，可刺激癌细胞的迁移和扩散；再如，参与糖酵解的3-磷酸-甘油醛脱氢酶（glyceraldehyde-3-phosphate dehydrogenase）以四聚体的形式存在于细胞质基质，而当以单体的形式存在于细胞核的时候，它却是一种尿嘧啶-DNA糖苷酶（uracil-DNA glycosylase），参与DNA的碱基切除修复（参看第三十四章“DNA损伤、修复和突变”）；还有大肠杆菌的生物素合成酶兼做其生物素操纵子的阻遏蛋白。

蛋白质在行使功能的时候，通常都会涉及到与配体（ligand）的结合，因此可根据配体的性质，来对蛋白质的功能进行分类。与蛋白质结合的配体可以简单地分为两类：第一类是配体在结合以后，化学结构会发生变化，变成了另外一种配体。例如，酶在催化反应的时候，作为底物的配体在与酶结合以后，受到酶的催化，变成了另一种配体——产物；第二类是配体在结合前后并没有发生化学变化，例如 Fe^{3+} 与转铁蛋白的结合，O_2 和血红蛋白的结合。

第二节　蛋白质结构与功能之间的关系

蛋白质的结构与功能之间的关系一直是生化学家和分子生物学家最关注的问题之一。经过对多种蛋白质的结构与功能关系的研究，科学家们已总结出了其中的一些基本规则。

一、蛋白质结构与功能关系的一般规则

（1）大多数蛋白质具有特定的三维结构，也具有特定的功能。一旦三维结构被破坏，蛋白质的功能随之丧失。少数蛋白质处于天然无折叠状态，但仍然具有功能（参看本节“（四）无折叠蛋白质的结构与功能”）。

（2）蛋白质的三维结构直接决定蛋白质的功能。

（3）蛋白质的一级结构决定其三维结构，因此也最终决定了蛋白质的功能。

（4）结构相似的蛋白质具有相似的功能。反过来，功能相似的蛋白质具有相似的结构，特别是三维结构。

（5）在不同物种体内功能相同的蛋白质具有相同和基本相同的三维结构，但一级结构是否有差异以及差异的程度往往取决于物种之间在进化上的亲缘关系。

以组成有氧生物呼吸链的关键成分——细胞色素c为例，这种蛋白质存在于所有的有氧生物体内，与原核生物的质膜和真核生物的线粒体内膜相联系，其功能是作为一种流动的电子传递体，往返于呼吸链的复合体III和IV，进行电子的传递（参看第二十章“生物氧化”），但真正传递电子的是与细胞色素c共价结合的血红素辅基上的铁离子。在对多种不同来源的细胞色素c的晶体结构进行研究后发现，它们在三维结构上都惊人地相似。而在对40种不同的真核生物的细胞色素c的一级结构进行比较分析后还发现，在构成细胞色素c的110个左右的氨基酸残基中，28个残基始终不变，这意味着这28个氨基酸残基是细胞色素c形成稳定的三维结构，或者是其行使正确的功能所必需的。进一步的研究结果还表明，在这28个高度保守的氨基酸中，有3个Gly、2个Cys、1个His和1个Lys。其中的3个Gly在所有的细胞色素c分子中都是绝对保守的；而2个Cys和1个His关系到血红素辅基与细胞色素c的共价及配位结合，因此也是不可变更的；1个Lys残基参与细胞色素c与膜的结合，因而也是高度保守的。从二级结构的层次来看，保守的氨基酸多分布在无规则卷曲中，还有5段β-突起高度保守。这与无规则卷曲通常与蛋白质的功能关系密切，以及β-突起通过细调肽链的走向而影响细胞色素c最终三维结构的形成是一致的。

Quiz1 与其他氨基酸残基相比，Gly在所有的蛋白质进化过程中是最保守的，为什么？

根据两种有氧生物在细胞色素 c 一级结构上的差异程度，可以判断它们之间的亲缘关系。例如，人与黑猩猩没有差别；而人与绵羊相差 10 个氨基酸残基，与鲫鱼相差 18 个，与酵母相差 44 个(图 3-1)。这清楚地表明，亲缘关系越近，氨基酸残基的差异就越少。根据不同种属之间氨基酸残基差异的多少和替换速率，可基本了解生物的进化过程，并描绘出系统分子进化树。

(6) 一级结构相似的蛋白质往往具有共同的起源。很多学者经常根据蛋白质一级结构的相似性来研究生物进化。一般说来，如果两种生物的亲缘关系紧密，那么它们的基因和蛋白质的一级结构就相似。

蛋白质进化研究中，以下概念比较重要：①类似物(analog)。专指具有相同的功能，但起源于不同的祖先基因的蛋白质，它们是基因趋同进化(convergent evolution)的产物。例如，牛和鼠疫杆菌(Yersinia pestis)都合成一种酪氨酸磷酸酶(tyrosine phosphatase)。这两种生物产生的同一种酶在活性中心的三维结构十分相似，活性也相似，但一级结构差别很大，显然它们是从完全不一样的祖先基因进化而来的。再如，枯草杆菌合成的一种丝氨酸蛋白酶，与哺乳动物体内的丝氨酸蛋白酶不仅一级结构不一样，三维结构也不一样，但活性中心都含有由 Ser、His 和 Asp 组成的催化三元体(参看第十章“酶的催化机理”)。显然，它们也来自不同的祖先基因。②同源物(homolog)。专指存在于不同生物或者同种生物，来源于某一共同祖先基因的蛋白质(图 3-2)。③种间同源物(ortholog)。也称为直向同源物或直系同源物，属于同源物中的一种，专指来自于不同物种的由垂直家系(物种形成)进化而来的蛋白质，它们通常保留与原始蛋白质相同的功能，但也不尽然。例如，小鼠、蛙和鸡各自的 α 珠蛋白或 β 珠蛋白(图 3-2)。④种内同源物或旁系同源物(paralog)。属于同源物中的另外一种，专指同一物种内由于基因复制、分离产生的同源物。例如，小鼠 α 珠蛋白和 β 珠蛋白，蛙的 α 珠蛋白和 β 珠蛋白，鸡的 α 珠蛋白和 β 珠蛋白(图 3-2)。通过进化，这一类种内的同源物可能会获得新的功能，但这种新功能多多少少会与原来的功能有一定的关系。

	黑猩猩	绵羊	响尾蛇	鲫鱼	蜗牛	烟草小菜蛾	面包酵母	花椰菜	欧防风
人	0	10	14	18	29	31	44	44	43
黑猩猩		10	14	18	29	31	44	44	43
绵羊			20	11	24	27	44	46	46
响尾蛇				26	28	33	47	45	43
鲫鱼					26	26	44	47	46
蜗牛						28	48	51	50
烟草小菜蛾							44	44	41
面包酵母								47	47
花椰菜									13

图 3-1 不同物种的细胞色素 c 一级结构的比较

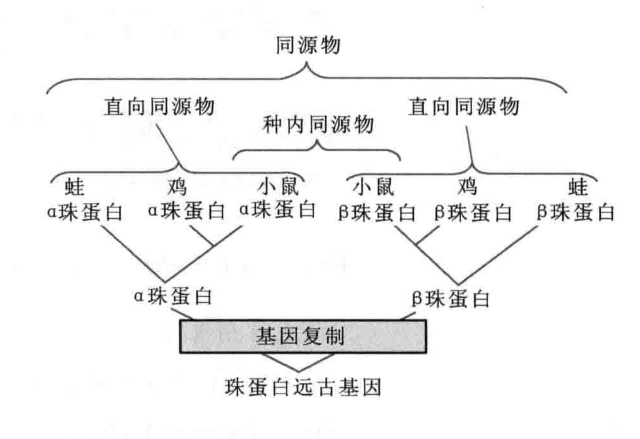


图 3-2 同源物、直向同源物和种内同源物之间的关系

现有多种专门的软件能够对以上四种情形进行分析，如 ClustalX，还有许多网站提供在线分析服务，如欧洲生物信息学中心(European Bioinformatics Institute，EBI)提供的 http://www.ebi.ac.uk/clustalw/，这为在分子水平上研究进化提供了便利。

(7) 许多疾病都是体内重要的蛋白质结构异常引起的(参看前一章的内容)。

二、几类重要的蛋白质的结构与功能

(一) 纤维状蛋白质的结构与功能

纤维状蛋白质倾向于形成有规则的纤维状结构，这种结构源于其一级结构的高度规律性——氨基酸残基的种类有限，但序列通常以重复的形式出现。纤维状蛋白质一般由几条肽链组成，难溶于水，化学反应性较差，如毛发中的角蛋白、弹性蛋白和胶原蛋白。其主要功能是在结构或机械支持上，为

Quiz2 你认为纤维状蛋白质通常缺乏何种二级结构?

机体提供支持和保护。

1. α 角蛋白

α 角蛋白来源于动物的毛发、角、蹄、喙和爪等部位。其一级结构由 311~314 个氨基酸残基组成。每一个 α 角蛋白分子在肽链的中央形成典型的 α 螺旋，而两端为非螺旋区。

螺旋区由 HXXHCXC 七肽重复序列组成，这使得两个 α 角蛋白分子能够通过 H 位的疏水 R 基团结合，相互缠绕形成双股的卷曲螺旋。双股卷曲螺旋大大地提高了 α 螺旋的稳定性。此外，在链间还形成了多个二硫键，这种共价交联可进一步提高 α 角蛋白的强度。指甲的强度比毛发高是因为含有更多的 Cys 残基，能形成更多的二硫键。

在美发过程中，无论是卷发还是直发，原理都一样：都是先用巯基化合物（如巯基乙酸胺或半胱氨酸）破坏二硫键（大约 45% 的二硫键被切断），使之被还原成自由的巯基，易于变形；再用发夹和发卷将头发塑成一定的形状（卷发使头发成为波状，直发则将头发拉直）；最后用氧化剂重建二硫键，使头发形状固定。

α 角蛋白分子还可以在双股卷曲螺旋的基础上，先形成原纤维，然后再由原纤维组装成纤维（图 3-3）。

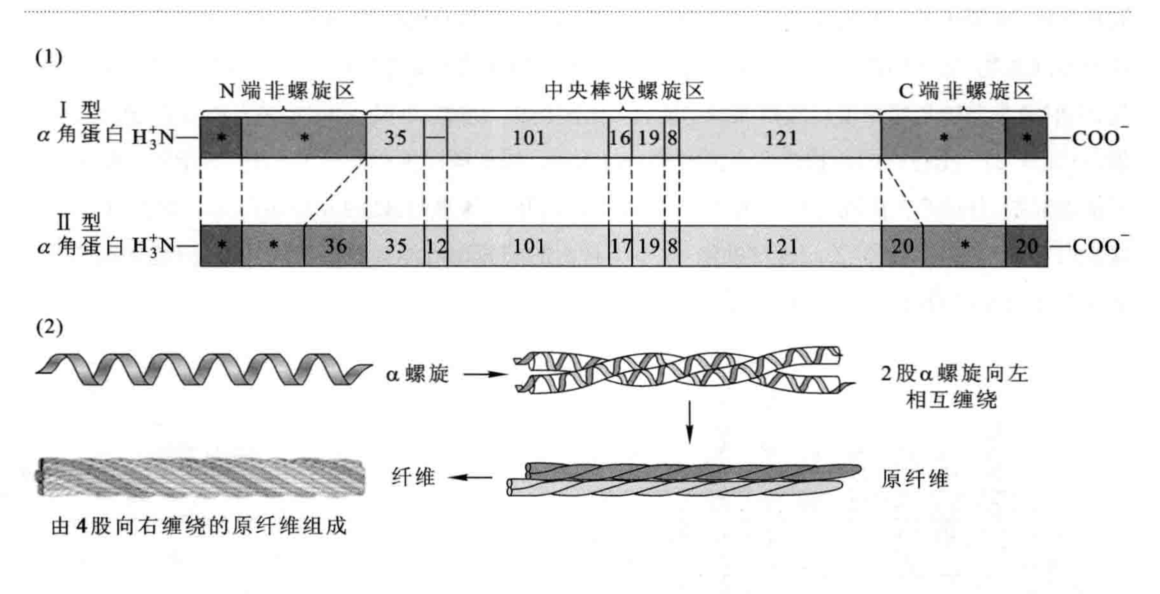


图 3-3 α 角蛋白的结构层次（Garrett 等，2010）

2. β 角蛋白

β 角蛋白主要来源于蚕丝和蜘蛛丝中的丝心蛋白，其一级结构富含 Ala 和 Gly，具有重复序列 Gly-Ala/Ser-Gly-Ala/Ser；它的二级结构主要是有序的反平行 β 折叠，还有一些环绕在 β 折叠周围的无规则卷曲和 α 螺旋（图 3-4）。有序的反平行折叠构成丝的微晶区（crystallite），由于 Gly 和 Ala/Ser 分别分布于折叠片层的两侧，相邻的 β 股能更加紧密地堆积形成网状结构（图 3-5），从而赋予丝较高的抗张性。而无序的 α 螺旋和无规则卷曲构成无定形区，又使丝具有一定的弹性。

3. 胶原蛋白

胶原蛋白作为动物细胞胞外基质内的一种主要结构蛋白，广泛存在于动物的结缔组织和其他纤维样组织中，如肌腱、韧带、骨骼和血管，是哺乳动物体内含量最丰富的蛋白质。食品工业上经常使用的明胶（gelatin）就是动物胶原蛋白经酸或碱部分水解的产物。

胶原蛋白的基本组成单位是由 3 条 α 链组成的原胶原（tropocollagen）。3 条 α 链可以相同，也可以不同。不同类型的原胶原由于 α 链的氨基酸组成及含糖量不同，因而性能也不同。4 个原胶原分子以平行交错的方式聚合成胶原原纤维，再进一步包装成胶原纤维。原胶原分子内部和原胶原单位之间会逐步形成特殊的共价交联，进一步稳定和加强胶原结构。

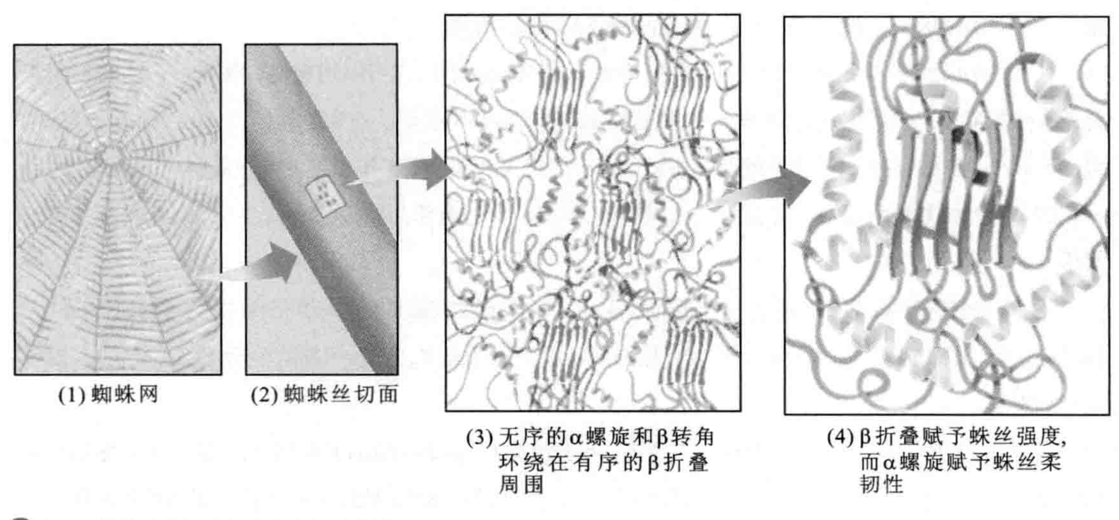


图3-4　蜘蛛丝中的丝心蛋白的结构层次

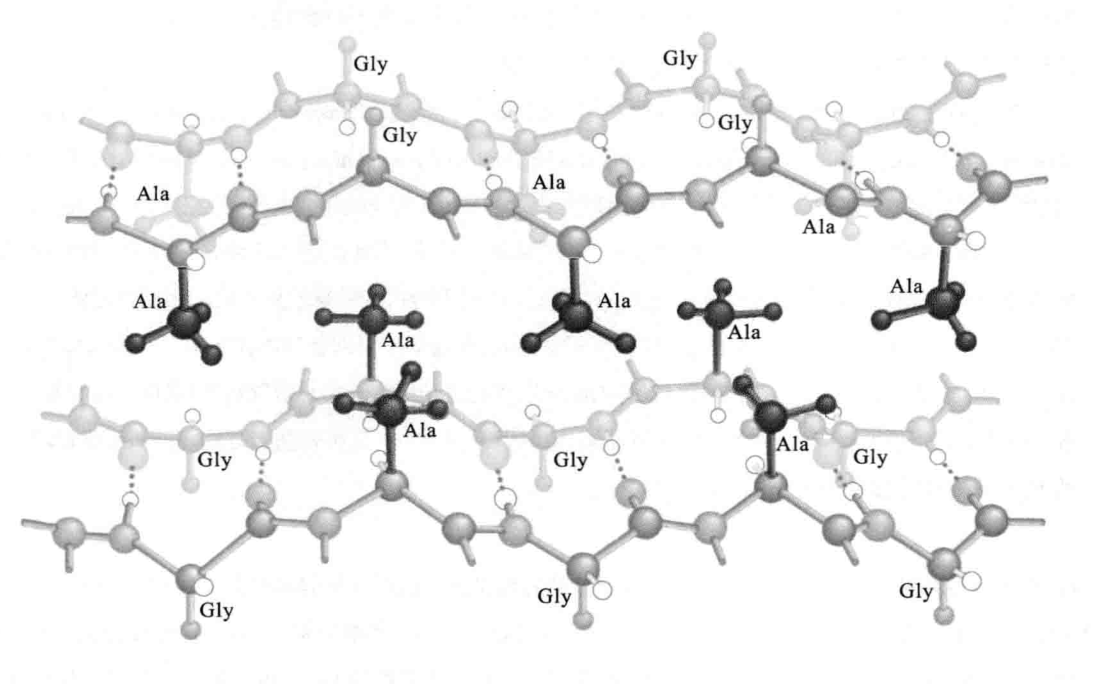
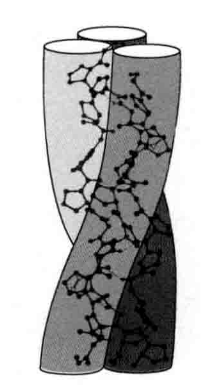


图3-5　β 角蛋白中的 β 折叠（Garrett 等，2010）



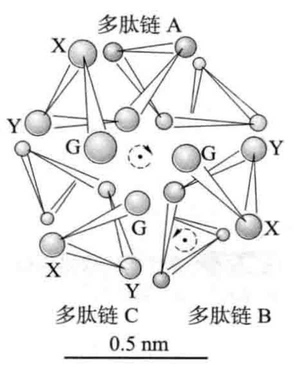


图3-6　胶原蛋白的三股螺旋

原胶原的一级结构的主要特征是：约 1/3 是 Gly（~33%），Pro 含量也很高（~12%），但 Tyr 含量少，Trp 和 Cys 缺乏；具有三种修饰的氨基酸，即 4- 羟脯氨酸、3- 羟脯氨酸（~9%）和 5- 羟赖氨酸；每一条肽链都具有重复的 Gly—X—Y 三联体序列（其肽链可用 $(\text{Gly—X—Y})_n$ 表示），重复次数在 200 左右。X 和 Y 通常是 Pro，也可能是 Lys；Y 位置上的 Pro 经常被羟基化为 4- 羟脯氨酸；Lys 也常被羟基化成 5- 羟赖氨酸，并且可能与相邻纤维上的正常 Lys 残基形成不寻常的共价交联。

胶原蛋白富含 Gly 和 Pro 的性质使得它难以形成 α 螺旋和 β 折叠，但有规律的三联体重复序列却有利于 3 条 α 链相互"抱成一团"，形成另外一种螺旋，即三股螺旋（triple helix，图 3-6）。三股螺旋为原胶原特有的二级结构，其二面角（Φ,ψ）为（$-51°$，$+153°$），由三股以左手螺旋存在的 α 链组成，这三股 α 链以氢键相连，并相互缠绕形成右手超螺旋。在螺旋中，体积最小的 Gly 正好位于螺旋的内部，构成紧密的疏水核心，而 Pro 和 Hyp 的侧链位于三股螺旋的表面，面向外，以尽量减少空间位阻。每一个 Gly 残基的 -NH 基团与 X 残基的 C═O 基团形成氢键，一个三联体序列大约形成一个氢键。三股螺旋比 α 螺旋更为伸展，每一个氨基酸残基上升 0.29 nm，一圈有 3.3 个氨基酸残基。

Pro 残基缺乏氢键供体，因此单凭 3 条肽链主链之间形成的氢键，还不足以稳定三股螺旋的结构。这

框 3-1 生化应用——蛛丝 / 蚕丝复合纤维的制备及其应用价值

作为一种天然的生物材料，蚕丝纤维除了作为伤口缝线以外，还可以用在伤口绷带、人工肌腱、组织骨架和微胶囊上。但同样是丝，蜘蛛丝的强度和弹性要明显高于蚕丝。有实验显示，一束由蜘蛛丝组成的绳子比同样粗细的不锈钢钢筋多承受 5 倍的质量！这相当于是一根铅笔粗细的蛛丝能阻止一架波音 747 飞机的飞行！蜘蛛丝纤维具有如此好的机械性能，使其非常适合作为精细的缝线，用于眼部、神经和美容等手术。

不过，蚕丝很容易通过养蚕业大量得到，但蜘蛛这种动物的领域性和同类相食的特性使得“养蛛业”行不通，因此难以大量制备蜘蛛丝。这就促使人们尝试使用低成本、方便、可靠的生物技术途径，去大规模生产蜘蛛丝。

沿着这个方向的一大突破首先是有人克隆了金丝蜘蛛（*Nephila clavipes*）蛛丝蛋白的 cDNA，包括壶腹状的丝心蛋白 1、丝心蛋白 2 以及长丝状的丝蛋白。当得到这些基因以后，人们立刻想到利用多种外源蛋白质表达系统（如大肠杆菌、酵母、昆虫细胞、植物细胞和哺乳动物细胞表达系统）去大量表达这些基因，但却遇到一个难以克服的问题，就是表达出来的蛋白质虽多，却很难像在蜘蛛丝腺内一样，绕成蛛丝纤维。这就使得人工大规模制备蛛丝纤维的想法一直难以实现。

然而，就在 2012 年新年伊始，美国科学院院刊上刊登了一篇题为“Silkworms transformed with chimeric silkworm/spidersilk genes spin composite silk fibers with improved mechanical properties”的论文，使这个问题似乎终于有了突破。来自美国圣母大学、怀俄明大学和我国浙江大学的研究人员研究出了一种方法，得到了一种蛛丝 / 蚕丝的复合纤维，其机械性能与天然的蛛丝差不多。他们的基本思路是：使用一种特殊的载体，将编码蛛丝蛋白的基因导入到家蚕的体内，培育出转基因家蚕，并使蛛丝蛋白的基因受蚕丝腺特异性启动子（DNA 分子上控制基因转录启动的特殊碱基序列，参看第三十六章“DNA 转录”）的驱动，在蚕丝腺内表达。表达出来的蛛丝蛋白能与蚕丝腺内原来表达的蚕丝蛋白一起，共同组装成蚕丝 / 蛛丝复合纤维。根据测定，这种复合丝纤维的性能比蚕丝强，与蛛丝不相上下。这种转基因蚕的问世，向大规模制备具有原始蛛丝特性的生物材料迈进了一大步。

时就需要通过特殊的化学修饰在肽链上引入额外的氢键供体。这种化学修饰就是发生在 Pro 残基上的羟基化反应。其中催化羟化反应的羟化酶（hydroxylase）需要 O_2、Fe^{2+} 和维生素 C。Fe^{2+} 包埋在羟化酶的活性中心，所起的作用是活化充当底物的 O_2，但它很容易被氧化成无活性的 Fe^{3+}。维生素 C 所起的作用是作为抗氧化剂，防止 Fe^{2+} 的氧化。故维生素 C 的缺乏会导致胶原的羟化反应不能充分进行，也就影响到正常胶原原纤维的形成。那些非羟化的前 α 链在细胞内很容易降解，从而导致牙龈出血、创伤不易愈合等病变。

胶原蛋白主要由成纤维细胞（fribroblast）合成。刚刚翻译出的多肽链被称为前 α 链，其两端各有一段不含 Gly—X—Y 重复序列的前肽。3 条前 α 链的 C 端前肽借助二硫键形成链间交联，使得 3 条前 α 链“对齐”排列，然后从 C 端向 N 端形成三股螺旋。前肽部分呈非螺旋卷曲，形成球状结构域。带有前肽的三股螺旋胶原分子被称为前胶原（procollagen）。前 α 链在细胞内合成后还要进行糖基化（glycosylation）修饰，才能自组装成三股螺旋。

Quiz3 变性后的胶原蛋白还能复性吗？为什么？

前胶原分泌到胞外以后，在前胶原肽酶（procollagenpeptidase）的催化下，两端的前肽序列被水解后成为原胶原。胶原变性后不能自然复性重新形成三股螺旋结构，是因为成熟胶原分子的肽链不含前肽，故不能再进行“对齐”排列。在细胞外基质中的赖氨酰氧化酶（lysyl oxidase）的催化下，胞外原胶原上的 Lys 残基被氧化成醛赖氨酸（allysine）。醛赖氨酸上的醛基可以与邻近肽链上的 Lys 的氨基或羟赖氨酸的羟基缩合，形成共价交联。共价交联的形成是一个缓慢的过程，可持续一生，故交联的程度随着年龄的增加而加深。引入共价交联能提高组织强度，但同时也降低了组织的弹性和柔韧性。

在人体内已发现的近 29 种不同类型的胶原蛋白中，Ⅰ型胶原最多，占 90% 以上。营养不良或者

基因缺陷可导致许多与胶原相关的疾病，如成骨不全(osteogenesis imperfecta)。成骨不全又称为脆骨病(brittle bone disease)，其典型症状为骨骼脆性增加。在患者体内，编码组成Ⅰ型胶原α链的基因有缺陷，而导致肽链上的某些 Gly 被侧链较大的氨基酸残基取代。较大的侧链基团容易产生空间位阻，这使得三股螺旋出现突起，破坏了螺旋的稳定性，而导致胶原蛋白不能正常地行使功能。

(二) 球状蛋白质的结构与功能

生物体内的大部分生物功能是依赖于球状蛋白质完成的。机体内有各种各样的球状蛋白质，这里只选择以两类与人体健康有密切关系的球状蛋白质——珠蛋白家族(globin family)和免疫球蛋白为例，详细分析它们的结构与功能之间的关系。

1. 珠蛋白家族

这一家族的蛋白质都含有血红素辅基，都能够可逆地结合氧气，都含有珠蛋白折叠这样的结构模体。属于这一类家族的蛋白质有：肌红蛋白(myoglobin，Mb)、血红蛋白(hemoglobin，Hb)、神经珠蛋白(neuroglobin，Ngb)和细胞珠蛋白(cytoglobin，Cygb)。

(1) 肌红蛋白　Mb 存在于肌肉中，在心肌中的含量特别丰富。其功能是为动物的肌肉组织贮备氧气，此外还能解除 NO 的毒性。水生哺乳动物(如鲸鱼)体内的 Mb 含量尤其丰富，因此它们可以在水下长时间憋气。

Mb 的一级结构特征包括：由一条肽链组成，含有 153 个氨基酸残基；紧密结合 1 分子血红素(heme)辅基。血红素由原卟啉(proporphyrins)和 Fe^{2+} 组成。

Mb 的二级结构特征包括：共有 8 段 α 螺旋，它们约占全部序列的 75%，按照从 N 端到 C 端的次序被依次编号为 A、B、C、D、E、F、G、H。螺旋之间是短的 β 转角或小环(CD 表示 C、D 螺旋之间的环，依此类推)(图 3-7)。有 4 个螺旋终止于 Pro 残基。

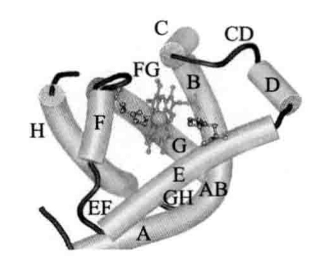


图 3-7　Mb 的三维结构以及其中各个螺旋的编号

Mb 是第一个获得完整三级结构的蛋白质。1959 年，John Kendrew 和 Max Perutz 使用 X 射线晶体衍射的方法，成功获得抹香鲸肌红蛋白的三级结构。1962 年，Kendrew 和 Perutz 由于测定了 Mb 及血红蛋白的三级结构而荣获诺贝尔化学奖。

Mb 整条肽链与血红蛋白的每一条肽链一样，折叠成紧密的球状结构，疏水侧链大都在分子内部，极性、带电荷的侧链则暴露在分子表面，因此水溶性好。其分子表面有一个深的疏水口袋，口袋的侧面由 E 螺旋和 F 螺旋组成，底部由 G 螺旋和 H 螺旋组成。血红素“藏”在袋中，与周围氨基酸残基形成次级键，而 Fe^{2+} 与 His F8 形成配位键(图 3-7)。该口袋允许 O_2 进入与 Fe^{2+} 结合，又可防止 Fe^{2+} 被氧化成 Fe^{3+}，但阻止 H_2O 的进入。Fe^{2+} 一共可以形成 6 个配位键，在没有结合氧气之前，它已形成了 5 个配位键(其中 4 个是与原卟啉吡咯环上的 N 原子结合形成，另 1 个是与 F 螺旋上的 His 残基(His F8)的咪唑基结合形成)。His F8 被称为近端组氨酸(proximal histidine)。显然第 6 个配位键是为 O_2 预备的，O_2 可以通过第 6 个配位键可逆地与 Fe^{2+} 结合(图 3-8)。但如果 Fe^{2+} 被氧化成 Fe^{3+}，水分子就会立刻占据第 6 个配位键，而导致氧气无法结合。CO 与 O_2 差不多大，因此也能与血红素结合。CO 的毒性是因为它与血红素的亲和力更大，从而阻止了 O_2 与血红素的结合。

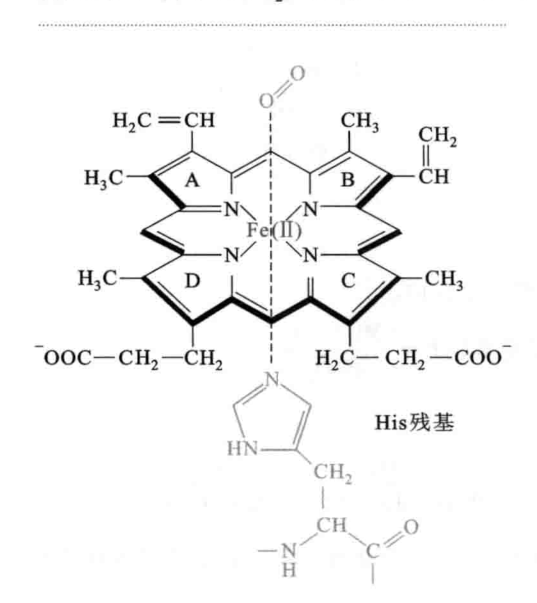


图 3-8　Mb 的血红素辅基

游离的血红素也能够与氧气结合，但它们在溶液中很容易相互靠近，因而导致在有氧的情况下，所有的铁最终都被氧化成高价态。而高价态的血红素铁是不能结合氧气的，因此生物没有选择用游离的血红素分子来运输或者贮存氧气。如果血红素结合在 Mb 或者

Hb 的疏水口袋之中，血红素分子等于被安全隔离起来，其中的铁也就难以氧化了。因此，Mb 的作用实际上是用疏水口袋保护血红素的二价铁，防止它被氧化。然而，尽管 Mb 和 Hb 本身已为上面的 Fe^{2+} 提供了很好的保护，但是细胞内的各种氧化剂仍然会不断地试图对其进行氧化，从而导致体内出现少量无功能的高铁血红蛋白（methemoglobin，met-Hb）和高铁肌红蛋白（metmyoglobin，met-Mb）。幸好正常人的体内还有一种保护机制，这种机制主要依赖于细胞色素 b_5 还原酶（cytochrome b_5-reductase）。该酶也被称为 NADH- 高铁血红蛋白还原酶（NADH-methemoglobin reductase），能够利用细胞内的还原性辅酶 I（NADH），将被氧化的铁还原成低价态。此外，有一种次要的 NADPH- 高铁血红蛋白还原酶也具有类似的功能。如果一个人编码细胞色素 b_5 还原酶的基因有缺陷，那他体内的 met-Hb 水平就可能异常高，而致患高铁血红蛋白症（methemoglobinemia）。由于大量 met-Hb 的存在而降低了红细胞的携氧能力，高铁血红蛋白症患者的血液颜色比正常人深，皮肤的颜色也因此可能呈现为罕见的蓝色。

将血红素放到 Mb 和 Hb 上还有一个好处，就是降低血红素与 CO 的亲和力。根据测定，游离血红素与 CO 的亲和力是与 O_2 亲和力的 25 000 倍！而 Mb 和 Hb 上的血红素辅基与 CO 的亲和力仅是与 O_2 亲和力的 200 倍。Mb 和 Hb 分子上的血红素辅基对 CO 亲和力的急剧下降与 E 螺旋内的第 7 号位 His 残基（His E7）有关。如图 3-9 所示，His E7 位于血红素平面的另一侧，与近端组氨酸（His F8）隔环相望，因此也称为远端组氨酸（distal histidine）。远端组氨酸和近端组氨酸一样，对珠蛋白家族的所有成员来说都是不可缺少的。远端组氨酸有两个重要的功能：一是保护血红素的二价铁，阻止细胞内任何可能的氧化剂对铁的氧化；二是为 CO 与血红素的结合制造障碍。因为对于 CO 而言，它与血红素铁结合的“舒适”角度是垂直于血红素平面（即 90°），而 His E7 的出现使得 CO 只能以 120°的角度结合，因此 CO 与血红素结合的亲和力就下降了。而对于 O_2 来说，它与血红素铁结合的“舒适”角度就是 120°，故有无 His E7 对血红素结合 O_2 没有影响。

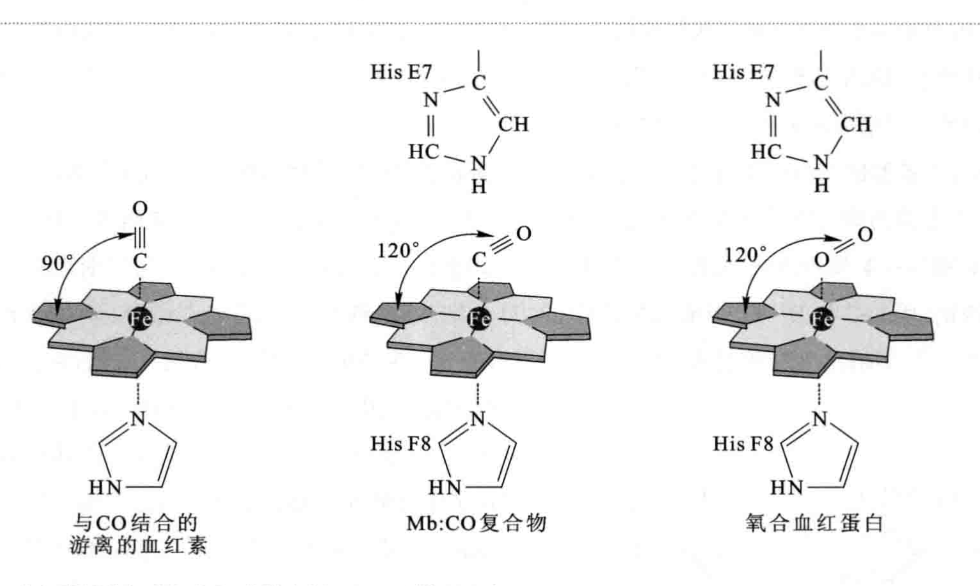


图 3-9　血红素辅基与 CO 或 O_2 的结合（Campbell 等，2009）

Mb 结合氧气的特征可以用氧合曲线来描述，为双曲线（图 3-10）。

Mb 的氧合反应可写成：$Mb+O_2 \longleftrightarrow MbO_2$，反应的解离常数 $K_d=\frac{[Mb][O_2]}{[MbO_2]}$

假定 Y 或 θ 为氧分数饱和度（fractional saturation），则：

$$Y=\frac{\text{被 } O_2 \text{ 结合的位点}}{\text{总的 } O_2 \text{ 结合位点}}=\frac{[MbO_2]}{[Mb]+[MbO_2]}=\frac{[Mb][O_2]}{K_d[Mb]+[Mb][O_2]}=\frac{[O_2]}{K_d+[O_2]}$$

由于溶解在液体内的气体浓度与液体上面的气体分压成正比，因此可用氧分压（p_{O_2}）代替氧气浓度。于是，$Y=\frac{p_{O_2}}{K_d+p_{O_2}}$。根据测定，在 Y=1 时，所有 Mb 上的氧气结合位点都被 O_2 占据。如果 Y=0.5，

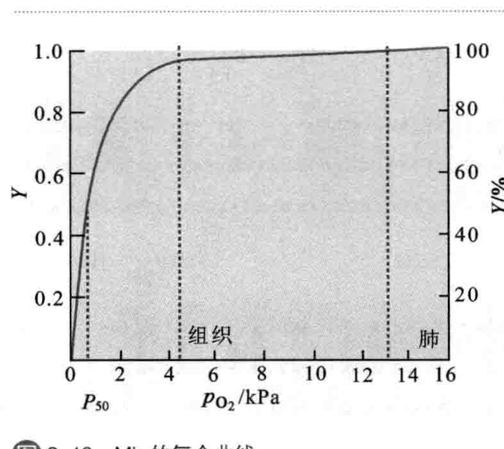


图3-10 Mb的氧合曲线

则$0.5=\frac{P_{50}}{K_d+P_{50}}$，即$K_d=P_{50}$。

从图3-10中可以看出，Mb倾向于结合氧气而不愿意放出氧气，因此它的功能是储存氧气。只有在p_{O_2}极低的时候，如肌肉因剧烈运动而缺氧，它才释放出氧气。

Mb在肌肉组织中的重要生理功能似乎表明它是机体不可缺少的蛋白质。然而，使用基因敲除（gene knockout）技术培育出的缺失Mb的小鼠能正常地运动和生育，对缺氧（hypoxia）也有正常的通气反应（ventilatory response）。尽管这些小鼠的心肌和比目鱼肌没有正常肌肉的颜色，但功能仍显示正常。缺失Mb的小鼠能够正常地生存，可能是因为机体为此做出各种代偿性反应，比如Hb浓度、血流量和毛细血管密度都有所提高。

(2) 血红蛋白　Hb存在于红细胞，其功能是运输氧气。它与Mb的比较参看表3-1。

表3-1　肌红蛋白和血红蛋白的比较

类别	肌红蛋白（Mb）	血红蛋白（Hb）
来源	肌肉细胞	红细胞
种类	一种	三种：HbA_1（成人98%）、HbA_2（成人2%）和HbF（胎儿）
一级结构	单条肽链，153个aa（氨基酸残基）	4条肽链，α亚基约含有141个aa，β亚基约含有146个aa，两者低于半数的氨基酸残基是相同的；α、β和Mb只有27个位置的氨基酸残基是相同的；HbA_1：$\alpha_2\beta_2$；HbA_2：$\alpha_2\delta_2$；HbF：$\alpha_2\gamma_2$
二级结构	75%为α螺旋，有A、B、C、D、E、F、G和H共8段螺旋，中间由无小环和β转角来连接	每条链同Mb，但D螺旋极短
三级结构	典型的球蛋白，内部含有珠蛋白折叠模体，分子表面有一个疏水口袋，血红素藏在其中	每条链同Mb
四级结构	无	4个亚基占据着四面体的4个角，链间以盐键结合，一条α链与一条β链形成二聚体；Hb可以看成是由2个二聚体组成的$(\alpha\beta)_2$，在二聚体内结合紧密，在二聚体之间结合疏松
辅基	血红素（Fe^{2+}），结合氧气	每个亚基结合1分子血红素（Fe^{2+}），1分子血红蛋白最多可结合4分子氧气
协同效应	无	正协同效应
Hill系数（h）	1	2.8
氧合曲线	双曲线	S形曲线
2,3-BPG	很难结合	两条β链之间可结合1分子BPG
Bohr效应	无	有
功能	在肌肉细胞中储存氧气	将氧气从肺部运输到外周组织

Hb由4个亚基组成，因而有四级结构。每一个亚基称为珠蛋白，单个亚基的一级结构与Mb差别较大，只有27个位置的氨基酸残基与Mb相同，但二级和三级结构却与Mb十分相似（图3-11）。

Hb氧合曲线为S形曲线（图3-12）。这意味着，只有在p_{O_2}很高的情况下（在肺部）Hb才结合氧气，而p_{O_2}一旦降低（在外周血管中），它就释放O_2，而此时的Mb却没有反应。就结合O_2的亲和力而言，4价的Hb还不如1价的Mb。Hb的氧合曲线形状与Mb不同是因为它与O_2的结合具有正协同效应（positive cooperativity）。

Hb的正协同效应是指Hb分子有一个亚基结合O_2后，其构象会发生变化，使得其他亚基对O_2的亲和力突然增强。协同效应可使用齐变或序变模型来解释（参看第九章“酶动力学”）。两种模型都

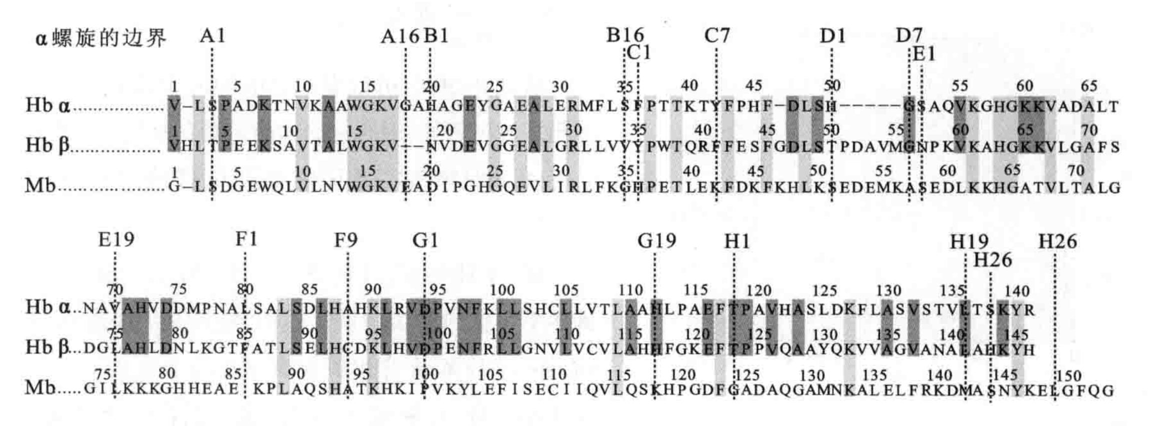


图 3–11 Mb 与 Hb 在一级结构上的比较（Nelson 等，2008）

假定血红蛋白存在两种构象，即紧张态（tense state，T 态）和松弛态（relaxed state，R 态）。在没有结合氧气时，Hb 的 4 条链之间结合紧密，这种构象称为 T 态，这种紧密结合是由盐键以及结合在 2 条 β 链之间缝隙中的 2,3– 二磷酸甘油酸（2,3–bisphosphoglycerate，2,3–BPG）造成的，它们屏蔽了分子表面疏水的空穴，使得 Hb 结合 O_2 的能力降低（游离的 α 链和 β 链结合氧气的能力与 Mb 相同）。

Quiz4 如果 Hb 被解离成单体，其氧合曲线还是 S 形吗？为什么？

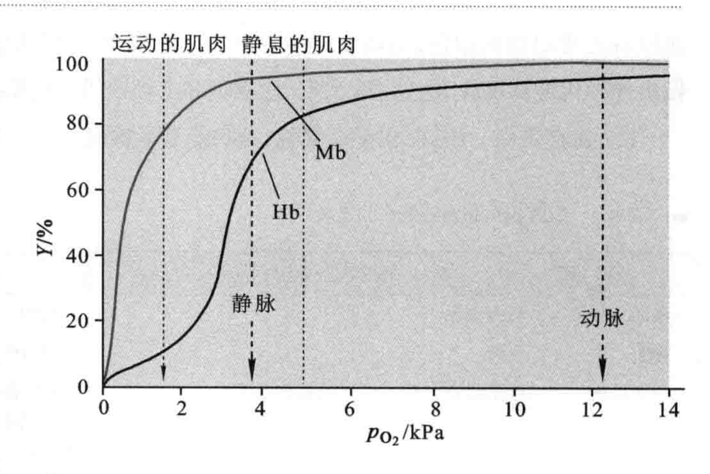


图 3–12 Mb 与 Hb 的氧合曲线

在脱氧状态下，Hb 上的 Fe^{2+} 由于邻近 His 残基和吡咯环 N 原子之间的空间位阻，而略偏离血红素平面（0.04 nm）。然而，一旦氧合，Fe^{2+} 就移向卟啉环，以便 O_2 能更紧密的结合。Fe^{2+} 的移位将近端 His 拉向血红素，近端 His 的移动又带动 F 螺旋也随之移动，而 F 螺旋的移动势必影响到它与相邻亚基的 C 螺旋之间的相互作用，最终导致相邻亚基的构象发生改变。这真可谓“牵一发而动全身”，于是相邻肽链之间的盐键遭到破坏，Hb 的四级结构也随之改变。这时 2 个二聚体（αβ）之间发生滑移，移动 15°，将 BPG 挤出。随后，四级结构发生进一步的变化，每条链表面疏水的空穴都暴露在外，这种构象称为 R 态，而这时 Hb 结合氧气的能力变强了（图 3–13，图 3–14）。

为了对 Hb 的正协同效应进行量化评估，需要在 Hb 的氧气分数饱和度方程中引入所谓的 Hill 系

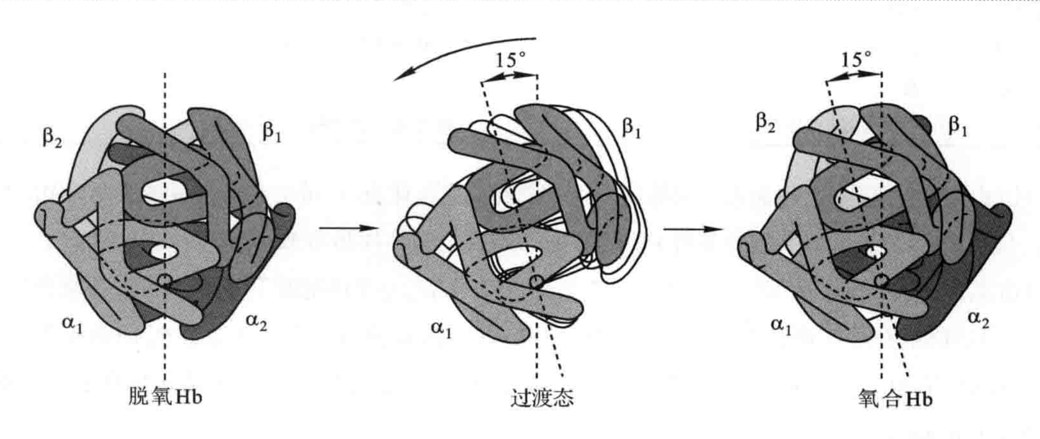


图 3–13 血红蛋白与氧气结合前后的构象变化

数(Hill coefficient,h)(参看第九章“酶的动力学”)。于是:

$$Y=\frac{[HbO_2]^h}{[Hb]^h+[HbO_2]^h}=\frac{(p_{O_2})^h}{(P_{50})^h+(p_{O_2})^h}$$

如果h=1,上面的方程实际上就是Mb的氧合方程,因此无协同效应;如果h>1,就有正协同效应;如果h<1,就有负协同效应。上面的方程可转换成:$\frac{Y}{1-Y}=\left(\frac{p_{O_2}}{P_{50}}\right)^h$

再进行线性化处理得:$\lg\frac{Y}{1-Y}=h\lg p_{O_2}-h\lg P_{50}$

如果以lg[Y/(1-Y)]为纵坐标,lg p_{O_2}为横坐标作图,这就是Hill作图。

根据测定,第一个亚基与O_2结合的P_{50}=4 kPa,最后一个亚基的P_{50}=0.04 kPa。由此可见,正协同效应导致Hb的最后一个亚基对氧气的亲和力增加了100倍。

除了正协同效应以外,还有另外两种效应也能影响到Hb的氧合能力。这两种效应分别是波尔效应(Bohr effect)和别构效应(allosteric effect)。

波尔效应是指H^+和CO_2促进Hb释放O_2的现象(图3-15)。波尔效应可解释Hb为什么在肺中吸氧排CO_2,而在肌肉中吸CO_2排氧。产生波尔效应的原因是H^+和CO_2能够与Hb特定位点结合,而促进Hb从R态转变为T态。与H^+引发的波尔效应相关的基团有:α亚基的N端氨基、α亚基的His122咪唑基以及β亚基的His146咪唑基。这三个基团在Hb处于T态的时候都是高度质子化的,而当氧气与Hb结合以后,质子发生解离。如果溶液中的pH降低,将有利于这三个基团处于质子化状态,从而稳定T态,抑制氧气的结合。用反应式来表示,即为:$Hb+O_2 \longleftrightarrow Hb(O_2)_4+nH^+$。显然pH下降,即$H^+$浓度升高,会使反应平衡向左移动,这时有利于Hb释放结合的氧气。

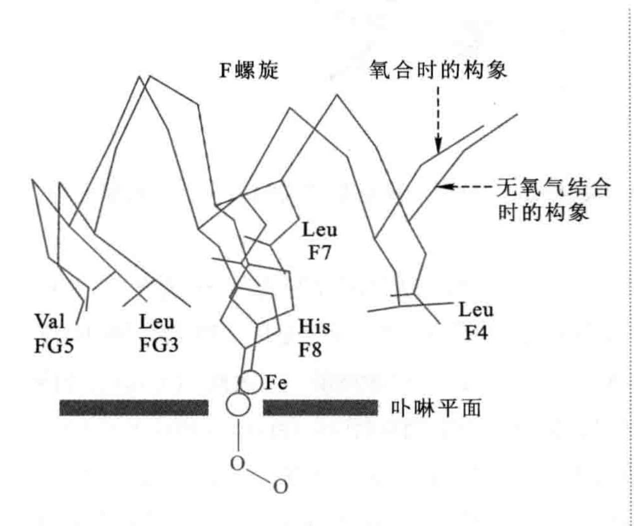


图3-14 血红蛋白与氧气结合前后F螺旋的构象变化

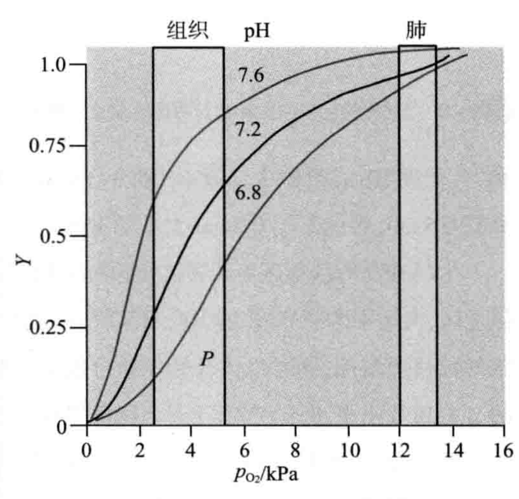


图3-15 质子浓度对Hb与O_2亲和力的影响

在外周组织,CO_2通过两种途径产生波尔效应。这两种途径不但能够促进O_2的释放,而且也有助于CO_2进入肺部呼出体外。第一种是在碳酸酐酶(carbonic anhydrase)的催化下,红细胞内的CO_2发生反应:$CO_2+H_2O \longleftrightarrow H_2CO_3 \longleftrightarrow H^++HCO_3^-$。此反应释放出质子而产生波尔效应,同时产生的$HCO_3^-$进入血浆,随循环到达肺部。第二种是$CO_2$与珠蛋白的N端氨基可逆地反应:$CO_2+Hb\text{-}NH_2 \longleftrightarrow H^++Hb\text{-}NH\text{-}COO^-$,形成氨基甲酸血红蛋白(carbaminohemoglobin),也释放出质子,从而产生波尔效应。CO_2主要通过这两种途径进入肺,分别占85%和15%,只有5%的CO_2以溶解的形式随循环进入肺。

在肺部,较高的p_{O_2}使得Hb能够有效地结合氧气,而正协同效应使得Hb从T态变成R态,并释放出质子。同时,与Hb氨基端结合的CO_2也被释放出来。释放出的H^+与随血液循环到达肺部的HCO_3^-在碳酸酐酶的催化下,发生逆反应,产生CO_2和H_2O。CO_2随后通过呼气排出体外。

别构效应是指除氧气以外的各种配体在血红素铁以外的位点与 Hb 结合，导致 Hb 的构象发生变化，进而影响到 Hb 氧合能力的现象。能与 Hb 结合的配体有 H^+、CO_2 和 2,3-BPG，它们统称为别构效应物（allosteric effector）。H^+ 和 CO_2 产生的效应就是波尔效应。2,3-BPG 产生的效应也是增强 Hb 在外周组织释放氧气的能力。

Quiz5 CO 与 Hb 的结合是否属于别构效应？为什么？

2,3-BPG 作为糖酵解的副产物广泛存在于红细胞（参看第二十二章"糖酵解"），其浓度与 Hb 不相上下，约为 5 mmol/L。2,3-BPG 只能与脱氧 Hb 内位于两条 β 亚基之间的带正电荷的空穴（1.1 nm）结合，稳定 T 态，显著降低 Hb 与 O_2 的亲和力，促进 Hb 在组织中释放 O_2（图 3-16，3-17）。氧合 Hb 位于两条 β 亚基之间的空穴已经明显变小，只有 0.5 nm，容纳不了 0.9 nm 大小的 BPG。

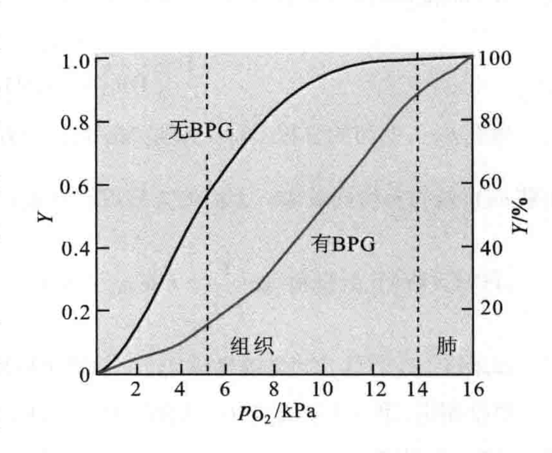


图 3-16 2,3-BPG 对血红蛋白与 O_2 亲和力的影响

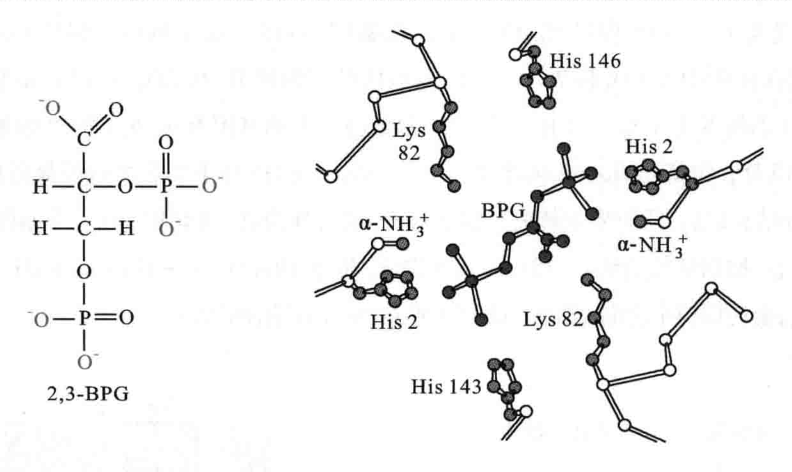


图 3-17 BPG 的化学结构及其与两条 β 链之间的结合

生活在高海拔地区可诱导红细胞内的 BPG 水平上升，这显然是生物对缺氧环境的一种适应。在高海拔地区，大气中的氧气较为稀薄，将氧气有效地释放到外周组织变得困难。为了适应这种环境，体内红细胞的数目会增加，同时机体开始加速合成 2,3-BPG（参看第二十二章"糖酵解"）。大概 24 h 以后，体内的 2,3-BPG 水平便开始持续上升。除此以外，贫血、肺功能衰弱和长期吸烟也可导致体内 BPG 水平上升。

Quiz6 试预测生活在高海拔的鸟类和低海拔的鸟类在 Hb 的结构与功能上会有何不同？

胎儿血红蛋白 HbF（$\alpha_2\gamma_2$）与 O_2 的亲和力明显高于成人的 HbA，这显然有利于胎儿通过胎盘（氧分压大大低于肺）从母亲那里获取 O_2。HbF 与 O_2 亲和力之所以高于 HbA，是因为与 α 亚基结合的 γ 亚基不能结合 BPG。由此可见，无私的母爱在分子水平上已经开始了。

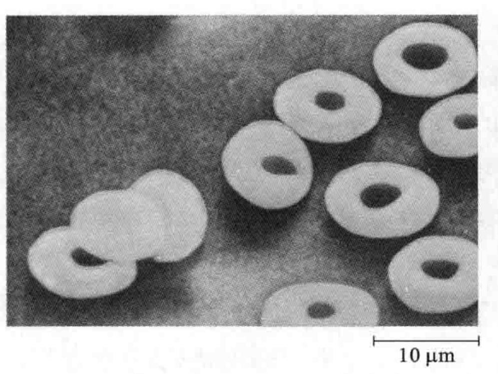
迄今为止，已发现多种 Hb 的突变体。根据表型，突变体可分为四种类型：① Hb 聚合成纤维状，致使红细胞成镰刀形；②改变与氧气的结合性质；③血红素辅基丢失；④四聚体解聚。

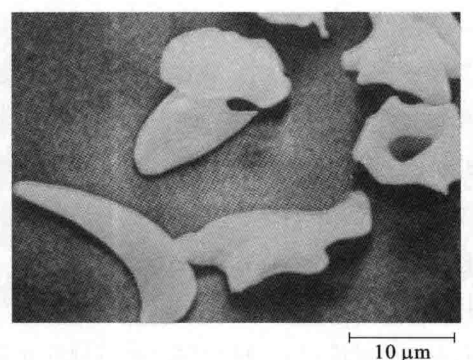
e3-1 Hb 突变体的结构及其功能异常

第一种突变体最为常见，它直接导致镰状细胞贫血症（sicklemia）。这种贫血患者的 Hb 简称为 HbS。HbS 与 Hb 在结合 O_2 的能力方面并没有区别，它们的区别在于 HbS 能造成红细胞溶血，致使病人体内的红细胞数量减少，通常只有正常人的 1/2。溶血后的 Hb 不能像红细胞中的 Hb 一样正常运输 O_2。病人表现为乏力，剧烈运动可导致死亡。HbS 导致溶血的原因在于其 β 亚基的 6 号位氨基酸残基从正常的 Glu 突变成 Val（图 3-18）。这种异常的 HbS 在脱氧状态下，相互间很容易通过表面的 Val 突起与疏水凹陷之间形成疏水键而聚集成纤维（图 3-19），导致细胞膜变形直至破裂。

Quiz7 若 βGlu6 → Asp6 或 Leu6 或 Ala，则变体的表型如何？

Ngb 和 Cygb 是近年来才发现的属于珠蛋白家族的两个新成员，它们与 Hb 和 Mb 一样也能结合氧





Val His Leu Thr Pro Glu Glu …
1 2 3 4 5 6 7
正常红细胞和正常
β亚基的一级结构

Val His Leu Thr Pro Val Glu …
1 2 3 4 5 6 7
镰刀形红细胞和异常
β亚基的一级结构

图 3-18 镰刀形红细胞和正常红细胞的外形比较以及 HbS 在一级结构上的变化

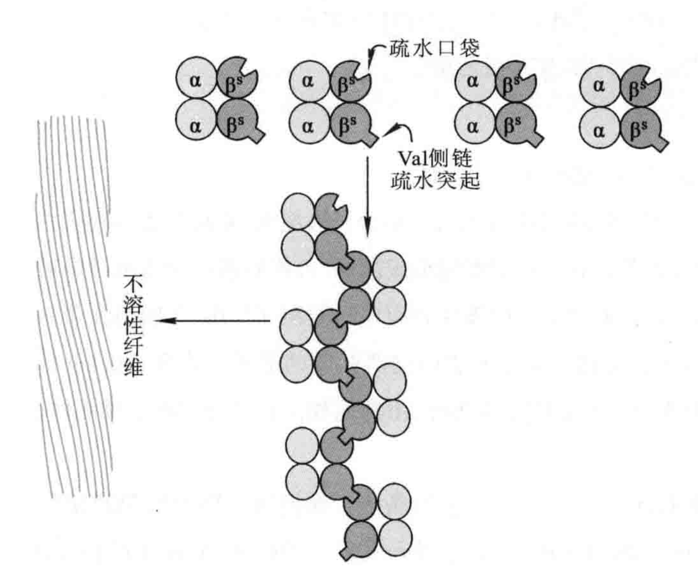


图 3-19 在脱氧状态下 HbS 之间的自组装反应（Baker，2001）

气。其中，Ngb 主要在脊椎动物的大脑和视网膜细胞中表达，而 Cygb 几乎在脊椎动物所有的细胞中都能表达。

e3-2 Ngb 和 Cygb 的结构与功能

2. 免疫球蛋白（immunoglobulin，IgG）

IgG 即抗体，是一类由 B 淋巴细胞分泌出的糖蛋白，可以与特异的抗原结合从而激发特定的体液免疫反应。

e3-3 IgG 的结构与功能

（三）膜蛋白的结构与功能

细胞内有 1/3~1/2 的蛋白质与膜结合或镶嵌在膜内，生物膜的各项功能主要是由各种结构不同的膜蛋白完成的。根据结构，膜蛋白可分为外周蛋白、内在蛋白和脂锚定蛋白。其中外周蛋白是可溶性的，与球状蛋白质非常相似，如细胞色素 c。内在蛋白的结构比较复杂，而脂锚定蛋白通过与其共价连接的疏水基团锚定在膜上（见第十六章“脂质与生物膜”）。

膜内在蛋白镶嵌于膜脂的特性，使这一类蛋白质经常位于细胞与外部环境的交界处，介导细胞与外界之间的信号转导，并执行很多重要的细胞功能。例如，它们作为各种神经递质、水溶性激素和其他配体的受体，构成各种跨膜的离子通道及转运蛋白等。但是，膜内在蛋白的疏水特性使其需要与生物膜共同形成稳定的天然构象，这大大增加了研究的难度。

与研究球状蛋白一样，要真正地理解膜蛋白的功能，首先需要确定它们的三维结构。然而，研究膜内在蛋白的三维结构并非易事，其困难集中反映在以下几个方面：

（1）含量低　例如一些跨膜受体蛋白的含量为微克级，而应用基因工程方法大量表达膜内在蛋白也存在着很大的困难。因此，迄今为止精细结构已被解析的主要是天然含量本来就高的膜内在蛋白，如牛心线粒体细胞色素 c 氧化酶、细胞色素 bc_1 复合体、植物光合作用的聚光复合体、紫色细菌光合反应中心、嗜盐菌菌紫质和细菌膜孔蛋白。

（2）分离、纯化比较困难　只有使用较剧烈的条件（如去垢剂、有机溶剂和超声波）才能将它们溶解下来，而分离后一旦除去去垢剂或有机溶剂，它们很容易聚合为不溶物。

（3）结晶困难　无论是电子晶体学方法还是 X 射线衍射方法，都需要膜内在蛋白的二维或三维晶

框 3-2　生化医药——丁酸的妙用

丁酸是一种常见的羧酸，现在人们发现它承载着有效治疗两种古老的人类遗传性疾病——镰状细胞贫血和β-地中海贫血的希望。镰状细胞贫血是由β珠蛋白发生一个点突变引起的，而β-地中海贫血的病因却是β珠蛋白不能正常表达。然而，这两种遗传性疾病并不影响胎儿，因为出生前和出生后几个星期，机体主要表达没有β珠蛋白的HbF，而不是HbA。HbF对氧气的亲和力高于HbA，这样可确保胎儿能通过胎盘从母亲的血液中获得足够的氧气。

用丁酸来治疗这两种遗传疾病的灵感来自两个偶然发现的交汇：一是有医生发现一些镰状细胞贫血患者只有轻微症状，原因是他们体内能继续产生高水平的HbF；二是一些糖尿病母亲所生的婴儿在出生以后很长一段时间内，继续产生大量的HbF。巧合的是，在这些婴儿的血液中，发现有高水平的氨基丁酸。这两个发现让美国奥克兰研究中心的Susan Perrine有了灵感，想到了通过"唤醒"在基因表达上处于"休眠状态"的胎儿珠蛋白基因来治疗镰状细胞贫血。于是，她和她的同事向3个镰状细胞贫血患者和3个β-地中海贫血患者注入了丁酸钠溶液。经过约三个星期的实验性治疗，这些患者体内的HbF水平上升高达45%。其中一个地中海贫血患者，所有的症状被完全逆转。而且，这种治疗方法的不良副作用很小。当然这种疗法只有在对大量患者进行更多的治疗以后才可以宣布完全成功。

体，因此晶体的生长成了目前解析膜内在蛋白三维结构最大的障碍。

德国生化学家Hartmut Michel在研究中巧妙应用可透析的两性小分子去垢剂，解决了膜内在蛋白难以结晶的难题。其原理是：当小分子的去垢剂达到一定浓度时形成微团，可将膜蛋白的疏水区屏蔽起来，这使膜内在蛋白在水中呈溶解状态。在去垢剂微团的极性表面之间的相互作用下，膜内在蛋白有可能形成结晶(图3-20)。1977年，Michel正是利用这种方法得到菌紫质的晶体。次年，Michel偶然发现，当把这种物质放在冰箱中时，可形成固态的玻璃状聚集体，由此他相信有可能获得三维晶体。幸运的是，他很快便获得了成功。

1981年，Michel利用分子筛层析技术首次得到了光合反应的晶体。他把这一研究成果投稿给*Nature*，但被拒绝。后转投向*The Journal of Molecular Biology*，才得到发表。1982年，Michel在Robert Huber领导的当时世界上最先进的X射线衍射实验室做报告时，引起了Huber的兴趣，二人选定Johann Deisenhofer完成紫色细菌光合反应中心的X射线晶体衍射分析。1985年，Deisenhofer完成了整个光合反应中心结构的测定。1988年，这三位杰出的德国生物化学家共享了诺贝尔化学奖。

对已获得三维结构的各种膜内在蛋白的研究表明，它们与非膜蛋白一样，含有相同的二级结构元件，但它们几乎完全存在于疏水环境之中，这种疏水环境是由膜脂上的疏水尾巴形成的。但由于肽键的亲水性质，将一条充分伸展的多肽链插入或贯穿于一个脂双层膜在能量上显然是不利的。于是，多肽链总是通过形成特殊的二级结构或三级结构来避免肽键与疏水的脂环境接触，以便插入到脂双层结构之中。实际上，疏水环境更有利于二级结构的形成，这是因为在疏水环境中没有水分子的竞争干扰，肽链内的氢键更容易形成。

就氨基酸分布而言，膜内在蛋白暴露在水相中的表面性质实际上与一般的球状蛋白相同，它们的差别主要在于与脂双层作用的氨基酸的性质和分布。如果不考虑与脂双层相互作用的表面，膜内在蛋白的内部与绝大多数球状蛋白差不多，但那些参与信号转导的离子通道或参与物质跨膜转运的孔蛋白除外。

要解决多肽链通过脂双层膜的能量不利的问题，最简单的方法是形成α螺旋，因为α螺旋满足了主链原子形成氢键的倾向，并使侧链基团面向脂环境。既然α螺旋的每一个残基延伸0.15 nm，那么一个跨膜的α螺旋约含有20个氨基酸残基，这相当于脂双层疏水部分的平均厚度，即2.5~3 nm(图3-21)。因此，一段由20个疏水氨基酸残基组成的α螺旋就足以横跨生物膜，这个数据实际上已成为

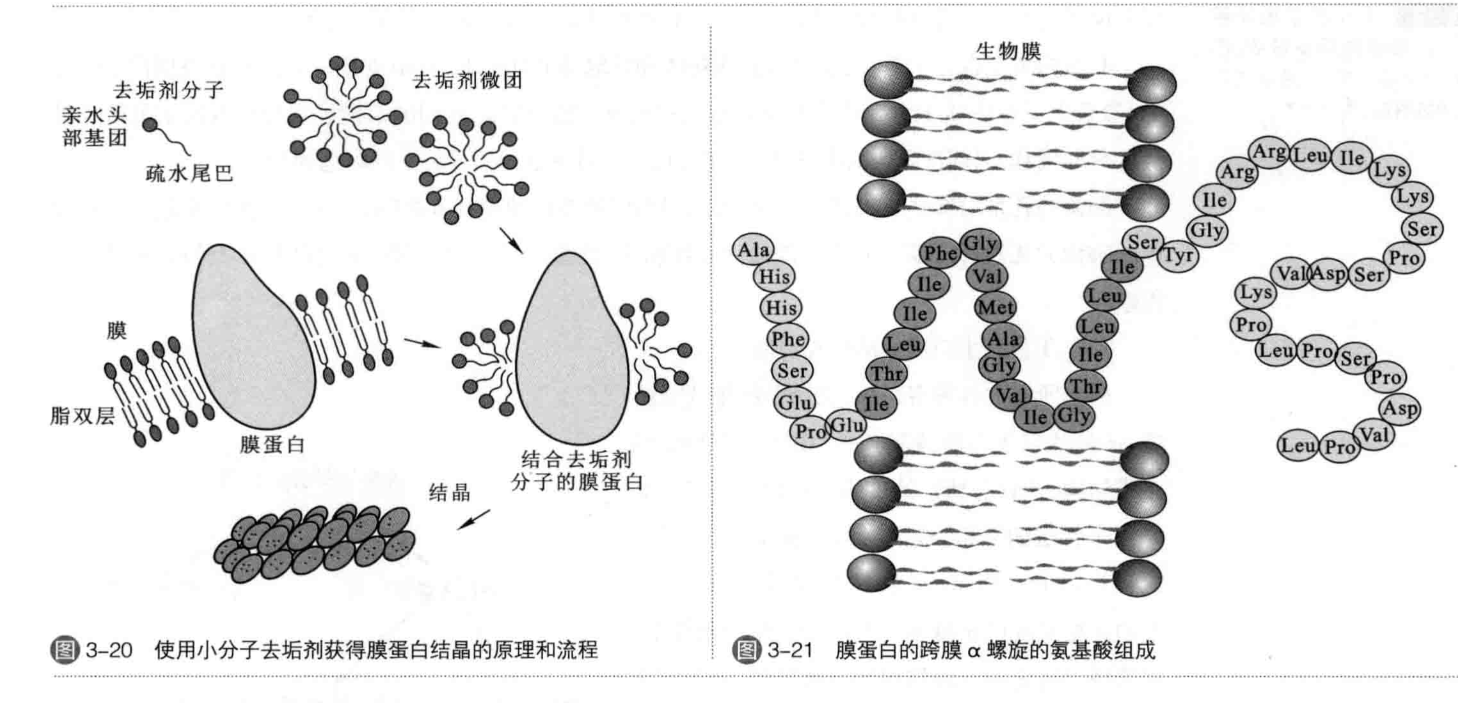


图 3-20 使用小分子去垢剂获得膜蛋白结晶的原理和流程

图 3-21 膜蛋白的跨膜 α 螺旋的氨基酸组成

鉴别一种蛋白质是不是膜内在蛋白的一个重要标志。

疏水作图(hydropathy plot)对于发现一个膜内在蛋白的跨膜螺旋十分有用,因为跨膜螺旋的疏水性一般要比通常的螺旋高得多。例如,在对羧酸转运感应蛋白(carboxylic acid transport sensor)进行疏水作图以后,预测其有两个跨膜螺旋,结果正好与实际相符(图 3-22)。

菌紫质是一种典型的膜内在蛋白。疏水作图表明,其含有 7 个跨膜螺旋,螺旋之间是相当短的小环(图 3-23)。其功能是作为一种受光子驱动的质子通道。受光子的激活,结合在菌紫质核心的视黄醛从 11- 顺式变成全反式。如此构型的变化会影响到一个氨基酸残基侧链的 pKa,从而导致 1 个质子从膜的一侧转运到另外一侧。

还有一种方法,可以解决多肽链通过脂双层膜遇到的能量不利的问题,那就是形成 β 折叠。然而,膜蛋白上的 β 折叠单凭序列是难以确定其存在的,原因是一个跨膜的 β 股只需要 8~9 个氨基酸残基就够了,但折叠一般由亲水氨基酸和疏水氨基酸残基交替组成。到目前为止,发现的全 β 折叠膜蛋白

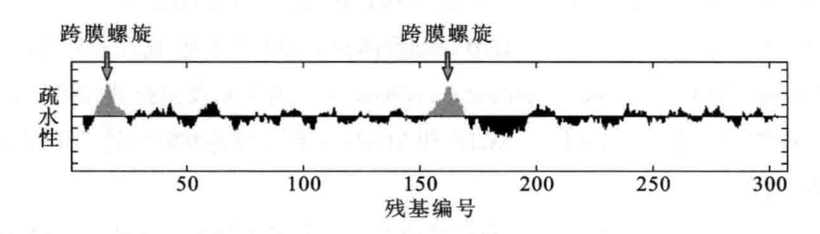


图 3-22 羧酸转运感应蛋白的疏水性作图

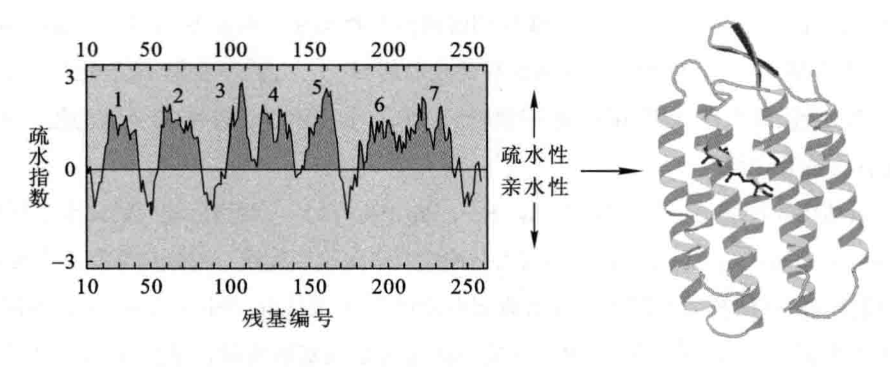


图 3-23 菌紫质的疏水作图结果以及实际的三维结构模型

Quiz8 蛋白质中也使用β－折叠横跨脂双层膜，然而，一个单一的β股从来不能单独跨膜，为什么？

具有反平行的β折叠桶结构，例如孔蛋白。联系相邻β股的是短的亲水性转角。

孔蛋白主要存在于革兰氏阴性菌、线粒体和叶绿体的外膜上，其中央是一个由β折叠桶构成的孔洞（参看上一章图2-17），允许小分子和水通过被动扩散跨膜。孔洞的两侧经常分布着带正电荷和带负电荷的残基，以创造跨膜的电场，使不同的孔蛋白对带电分子具有不同的选择性。

细菌的孔蛋白都是三聚体，它有很大的表面区域被包埋在亚基之间，这与球状蛋白亚基之间被包埋的表面并无两样。面向脂双层的表面高度疏水，比蛋白质的内部和亚基之间界面上的疏水性要高得多。

（四）无折叠蛋白质的结构与功能

前面所述的各种蛋白质，无论是纤维状蛋白质，还是球状蛋白质或膜蛋白，由于具有特定的三维结构，才具有特定的生物学功能。一旦它们失去了特定的结构，成为松散无折叠状态，就立刻丧失功能。然而，自20世纪90年代初以来，人们开始发现越来越多的蛋白质在生理条件下，尽管缺乏特定的二级结构和三级结构，处于完全无折叠或部分无折叠状态，但仍然具有功能。这类蛋白质现在统称为天然无折叠蛋白质（naturally unfolded protein，NUP），或固有无结构蛋白质（intrinsically unstructured protein，IUP），或固有无序化蛋白质（intrinsically disordered protein）。

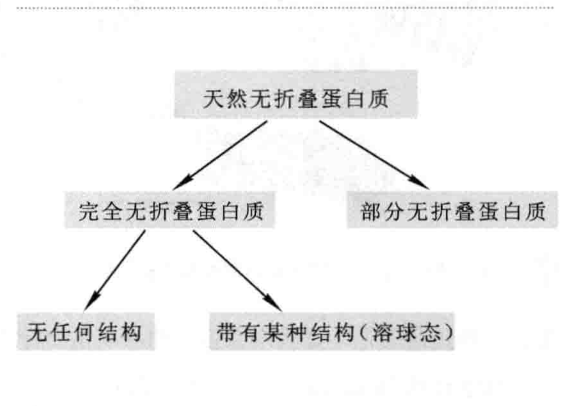


图3-24 天然无折叠蛋白质的分类

迄今为止，已发现的NUP约占蛋白质总数的30%。NUP一般可分为两类（图3-24）：一类是完全无折叠蛋白质（fully unfolded protein），约占10%；另外一类是部分无折叠蛋白质（partially unfolded protein），一般含有一段由多于50个氨基酸残基组成的无折叠区域。完全无折叠蛋白质还可以进一步分为两个亚类：第一亚类无任何二级结构；第二亚类无三级结构，与天然折叠蛋白质在折叠过程中形成的"溶球态"中间体相似。显然，天然无折叠蛋白质的发现是对传统的蛋白质结构与功能关系基本法则的挑战。



无任何结构的ACTR蛋白

+



溶球态的NCBD

相遇结合
发生折叠



ACTR－NCBD复合物

图3-25 两种NUP的相互作用导致各自发生折叠

在天然无折叠蛋白质中，有些必须处在无折叠状态才具有一定的功能，有些需要跟特定的配体（经常是DNA）结合，进而发生折叠，再行使功能。例如，淋巴类增强子结合因子1（lymphoid enhancer-binding factor-1，LEF-1）就属于后一类，它在B细胞前体和T细胞中表达，定位在细胞核，与DNA分子上的T细胞受体-α增强子（T-cell receptor-alpha enhancer）结合以后发生折叠，然后刺激相关基因的表达。再如，真核细胞中有两类转录因子ACTR和NCBD，它们单独存在时均处于无折叠状态，一旦结合，则立刻折叠（图3-25）。

NUP的主要功能是参与信号转导、细胞周期调控和基因表达调控。此外，它与翻译后加工也有关系，还经常充当RNA和蛋白质的分子伴侣。在信号转导过程中，经常涉及可逆的蛋白质磷酸化（reversible protein phosphorylation），这种磷酸化和脱磷酸化修饰是真核细胞调节蛋白质或酶活性的主要手段之一（参看第十七章"激素及其受体介导的信号转导"）。人们还发现，磷酸化位点周围的氨基酸组成、序列复杂性、疏水性、电荷和其他序列特征与NUP十分相似。因此，有人认为蛋白质磷酸化位点主要发生在无折叠区域。

NUP一级结构的特征是含有较多的Gln、Ser、Pro、Glu和Lys，而侧链较大的疏水氨基酸很少，如Val、Leu、Ile、Met、Phe、Trp、Tyr，因此可以利用上述性质来鉴别或预测一种蛋白质是不是无折叠蛋白。现在有专门的网站，在线提供预测天然无折叠蛋白质的服务，如http://bip.weizmann.ac.il/fldbin/findex。NUP为什么不能折叠或者只能部分折叠？主要原因是疏水氨基酸残基含量少，而疏水作用力是驱动蛋白质折叠的主要动力。

Quiz9 为什么NUP中侧链较大的疏水氨基酸残基含量少？

NUP 虽然处于完全无折叠或部分无折叠状态，但并不比处于折叠状态的蛋白质更容易水解，这主要是因为 NUP 通常分布在细胞核和无蛋白酶的细胞器，远离有蛋白酶的区域，而且多数 NUP 在遇到合适的配体以后会发生折叠。

从生物进化学的角度看，NUP 似乎具有很多优势，可能包括：①不需要经过折叠就能起作用；②可以结合几种不同的配体，而表现多种功能；③拥有较大的分子之间相互作用的界面，有利于分子识别；④有利于细胞信号转导过程中开与关的切换。这些优势对于复杂的生物来说更有用，因此 NUP 的分布与生物复杂性正相关，即生物越复杂，NUP 越多。NUP 在真核生物中比较普遍，在原核生物中则少见。

第三节　蛋白质的功能预测

就目前的研究水平，一般很容易从蛋白质的序列和结构推导出其生化功能，有时还可能研究得更深。在当今基因组学和后基因组学时代，通常凭经验，再综合各种不同的技术进行功能预测。尽管预测一种蛋白质功能最好的方法是先得到它的三维结构，但也可以单独进行功能预测或者与结构预测结合起来。

目前经常使用的功能预测法主要包括：

1. 基于序列的途径（sequence-based approach）

如果蛋白质 A 有功能 X，而蛋白质 B 是蛋白质 A 的同源物，则 B 就有功能 X。

序列比对是检测同源基因的强有力工具，但限于亲缘关系不太远的基因组。

2. 基于结构的途径（structure-based approach）

假定蛋白质 A 有功能 X，而且具有某种结构特征，那么 A 的功能位点便具有某种结构特征。如果蛋白质 B 具有这种结构特征，蛋白质 B 就可能具有功能 X。

这种基于结构的方法可能检测出基于序列途径不能确定的遥远的同源物。

3. 结构与序列比对法

除了使用序列信息以外，还使用结构信息。

蛋白质结构预测线索法（protein threading）是一种很流行的方法。这种方法首先假设多肽的构象，然后根据得到的结构计算出其能量。通过计算各种已知结构的能量，可以得出与给定蛋白质序列最符合的构象。由于该结构是假设的而不是计算出来的，线索法有时指的就是“反向蛋白质折叠”。

SCOP（http://scop.berkeley.edu/）是一个在线的蛋白质分类数据库，它根据蛋白质的结构和功能的相似性，将蛋白质分成若干个等级，按照从低到高的排列依次是：家族（family）、超家族（super-family）、栏（fold）和类（class）。其中类是最高级水平，共分为 11 类，包括全 α 蛋白（all-alpha protein）、膜蛋白（membrane protein）、αβ 蛋白（alpha-beta protein）、螺旋卷曲蛋白（coiled-coil protein）等；栏为第二级水平，共有 800 个栏；超家族位于第三级，共有 1 294 个超家族；家族处于第四级，共有 2 327 个家族。

使用序列 – 结构比对法，可以预测一种蛋白质属于哪一个家族、超家族或栏。被预测的处于同一个 SCOP 家族的蛋白质被认为是直向同源物；处于同一个 SCOP 超家族的蛋白质被认为是同源物；处于同一栏的蛋白质被认为是类似物。

通过这种方法，还可以对一种蛋白质的配体结合位点或大分子结合位点进行预测：已知约 85% 的配体结合位点是最大的裂缝，还有约 10% 的配体结合位点是第二大裂缝；大分子结合位点（蛋白质、DNA 和 RNA）和无序区域之间有密切的关系，而蛋白质序列中属于无序区的残基可通过计算的方法预测出来。

4. 基于模体的方法（motif-based approach）

假定一组蛋白质有功能 X,并且它们都有模体 Y,而蛋白质 A 具有模体 Y,那么蛋白质 A 的功能就可能与 X 有关。这种方法依赖于已鉴别的序列模体。Prosite(http://expasy.org/prosite)含有上千个蛋白质家族特异性的序列模体。而 ScanProsite 允许扫描一个蛋白质序列,从而发现存储在 Prosite 中的模体和资料(profile)。

Prosite 数据库是第一个蛋白质序列二次数据库,于 90 年代初期开始构建,现由瑞士生物信息学研究所 SIB 维护。Prosite 数据库是基于对蛋白质家族中同源序列多重序列比对得到的保守性区域,这样的区域通常与生物学功能有关,例如酶的活性位点、配体或金属结合位点等。因此,Prosite 数据库实际上是蛋白质序列功能位点数据库。通过对 Prosite 数据库的搜索,可判断该序列包含什么样的功能位点,从而推测其可能属于哪一个蛋白质家族。

5. 基于"连坐"(guilt-by-association)的功能预测

假定蛋白质 A 有功能 X,蛋白质 B 经常与蛋白质 A 相联系,那 B 的功能就可能与 X 相关。基于连坐的功能预测是一种非同源性途径。

科学故事——一种多功能蛋白质的发现

以前分子生物学里有一个基本观点,就是一个基因决定一条多肽链,而一条多肽链只具有一种特殊的功能。但随着分子克隆技术的发展,越来越多已知活性的蛋白质及其基因被克隆、测序和比对。在对许多蛋白质进行结构分析后发现,蛋白质结构与活性之间已不再是单纯的"一对一"的关系了。1987 年,芬兰 Oulu 大学的 Pihlajaniemi 在研究人体胶原蛋白的时候,就发现了这样的例外。

胶原蛋白是人体中含量最多的蛋白质,其二级结构是由 3 条相同的多肽链互相缠绕而形成的三股螺旋,每条肽链上含有多个 Gly—X—Y 重复序列,其中 X 位置主要是 Pro,Y 位置主要是羟脯氨酸(Hyp)。

胶原蛋白三股螺旋结构的稳定性与 Y 位置上有多少脯氨酸被羟基化有关,因此,催化 Pro 羟基化的脯氨酸羟化酶(proly1-4 hydroxylase)是决定胶原蛋白稳定性的一个十分重要的酶。脊椎动物的这种酶是由 2 个 α 亚基和 2 个 β 亚基组成。α 亚基和 β 亚基分别由不同的基因编码。不论 α 或 β 亚基,单独都没有活性,但 β 亚基的单克隆抗体能使整个酶失去活性,而只有 α 和 β 亚基同时存在的时候,酶才能与它的辅助因子——抗坏血酸结合,这说明 β 亚基的确是表现酶活性的重要组成部分。但奇怪的是,α 及 β 亚基在细胞内的表达并不一致。β 亚基的数量远远高于 α 亚基,因此,α 亚基合成后立刻会与 β 亚基结合,形成有活性的脯氨酸羟化酶,而留下许多单独存在的 β 亚基。那么单独的 β 亚基会有其他的功能吗?

既然蛋白质的一级结构决定蛋白质的高级结构,而蛋白质的高级结构又决定它的功能,那么如果知道了一种蛋白质的一级结构,就不仅可以预测和解释这种蛋白质的高级结构和功能,同时还可以将这种蛋白与其他已知功能的蛋白质进行序列比对,从而有可能发现新的功能。Pihlajaniemi 在得到脯氨酸羟化酶 β 亚基的一级结构以后,便将它的氨基酸序列输入计算机,在互联网上和其他已知蛋白质的氨基酸序列进行比对。出乎意料的是,她发现人类脯氨酸羟化酶 β 亚基的氨基酸序列有 94% 与小鼠的蛋白质二硫化物异构酶(PDI)相同。她用纯化的 β 亚基也检测到它的确有 PDI 的酶活性,再用克隆的 β 亚基的 cDNA 作探针,发现人类细胞中只有一个基因具有类似的碱基序列。因此 PDI 很可能就是单独的脯氨酸羟化酶的 β 亚基。换句话说,一个基因表达出来的蛋白质在不同的条件下可以同时具有两种完全不同的酶活性。

早在 1963 年 PDI 就在鼠肝和胰腺组织中被发现,其主要功能是催化蛋白质上二硫键的交换反应,具有较广的底物特异性。PDI 在体外能够促进被变性和还原的含有二硫键的蛋白质重获活性。当一个蛋白质的天然构象遭尿素破坏,而其二硫键也被还原剂打破以后,若要恢复这个蛋白质原来正确的构象,不仅要除去尿素,同时还要让二硫键重新形成,并恢复到原来正确的连接,PDI 就可以加速这个"尝试错误、发现正确"的过程。但 PDI 在体内是否也扮演同样的角色则仍是未知数。有了这个后知之明,再去看 PDI 在体内分布的情形,发现 PDI 的活性主要集中在内质网,这很容易让人们想起它可能在分泌蛋白折叠过

程中起作用，因为许多分泌蛋白都含有二硫键。而后来的许多研究证明，PDI 的确就具有这种功能。

那么，β 亚基的功能是不是就到此为止了呢？ 1991 年，Wetteran J.R. 等人使用氨基酸序列分析和免疫化学分析，发现了一种由大、小两个亚基组成的微粒体甘油三酯转移蛋白（microsmal triglyceride transfer protein，MTP）的小亚基与脯氨酸羟化酶的 β 亚基一模一样。但 MTP 的 PDI 酶活性仅是自由 PDI 的十分之一。后来又有人发现，单独的 β 亚基在细胞内还充当甲状腺素结合蛋白（thyroid hormone binding protein，THP）。

迄今为止，已经发现了许多多功能蛋白质，像哺乳动物的脂肪酸合酶居然具有 7 种酶活性，外加一种酰基载体蛋白的功能。多功能蛋白质的存在不仅可以提高基因的编码功能，还可能有利于对蛋白质的活性进行调控。

e3-4 本章小结填填看

推荐网址：

1. http://www.rcsb.org/pdb/101/motm.do?momID=41（内有 Hb 的 PDB 格式的三维结构，有助于理解 Hb 的结构与功能）
2. http://themedicalbiochemistrypage.org/hemoglobin-myoglobin.html（医学生物化学课程网站中有关 Mb 和 Hb 的内容）
3. http://www.sbc.su.se/~maccallr/thesis/node5.html（隶属于瑞典斯德哥尔摩大学生物信息学中心，有关蛋白质结构与功能的内容）
4. http://en.wikipedia.org/wiki/Membrane_protein（维基百科关于膜蛋白的内容）

参考文献：

1. Singh V, et al. Isozymes, moonlighting proteins and promiscous enzymes *Current Science*, 100(8): 1 152–1 160, 2011.
2. Hankeln T, et al. Neuroglobin and cytoglobin in search of their role in the vertebrate globin family. *J Inorg Biochem*, 99: 110–119, 2005.
3. Peter T. The interplay between structure and function in intrinsically unstructured proteins. *FESB Letters*, 15: 3 346–3 354, 2005.
4. Weissmann C, Flechsig E. PrP knock-out and PrP transgenic mice in prion research. *Br Med Bull*, 66: 43–60, 2003.
5. Jeffery C J. Moonlighting proteins: old proteins learning new tricks. *Trends Genet*, 19: 415–417, 2003.

第四章　蛋白质的性质、分类及研究方法

蛋白质是由氨基酸组成的多聚物,因此所具有的很多性质是由组成它的氨基酸残基带来的,例如紫外吸收、两性解离和颜色反应等。但作为生物大分子的蛋白质又有许多特有的性质,例如变性、复性和水解等。蛋白质的特有性质和蛋白质之间在性质上的差异是建立蛋白质分离和纯化方法的基础。

本章将主要介绍蛋白质的理化性质,以及如何利用这些性质对蛋白质进行分离、纯化,并对蛋白质的分类以及一些研究方法作简单的介绍。

第一节　蛋白质的理化性质

蛋白质的理化性质主要包括紫外吸收、两性解离、胶体性质、沉淀反应、变性、复性、水解和颜色反应等。

1. 紫外吸收

Trp、Tyr 和 Phe 三种芳香族氨基酸的 R 基团在 280 nm 波长附近有最大的紫外吸收,绝大多数蛋白质都含有它们,因而也会有紫外吸收现象。核酸也有紫外吸收,但最大吸收峰在 260 nm。因此,测定 280 nm 的光吸收是对蛋白质进行定性和定量分析较简捷的方法。

2. 两性解离

与氨基酸、寡肽和多肽一样,蛋白质也能发生两性解离,具有 p*I*。蛋白质的两性解离性质是由其表面氨基酸残基可解离的 R 基团以及肽链两端游离的氨基或羧基造成的。由于一个蛋白质分子含有多个氨基酸残基,其解离情况要比单个氨基酸或一个小肽复杂得多,并且各个可解离基团的 pKa 究竟是多少很难确定,因此一种蛋白质的 p*I* 难以直接计算,一般使用等电聚焦或等电点沉淀等方法进行测定。

不同蛋白质的氨基酸组成不同,因此 p*I* 会不一样。p*I* 不同的蛋白质在同一 pH 条件下所带净电荷不同,因而可用离子交换层析或电泳的方法对它们进行分离和纯化。如果蛋白质的碱性氨基酸残基较多,则 p*I* 偏碱;如果酸性氨基酸残基较多,p*I* 偏酸。而酸、碱氨基酸含量相近的蛋白质的 p*I* 大多为中性偏酸。

3. 胶体性质

可溶性蛋白质在水溶液中具有电泳、布朗运动、丁达尔现象和不能通过半透膜等典型的胶体性质。

蛋白质之所以能以稳定的胶体形式存在,是因为:①蛋白质分子大小已达到胶体质点范围,具有较大的比表面积。②蛋白质分子表面有许多亲水基团,这些基团能够与水分子形成氢键,因而对水分子有高度的亲和性,很容易吸附水分子,形成水化膜。水化膜的存在使得蛋白质颗粒彼此难以靠近,增加了蛋白质在溶液中的稳定性,防止它们从溶液中聚集或沉淀出来。③同一种蛋白质分子在非等电状态时带有同性电荷而相互排斥,不会聚集沉淀。当然,如果这些稳定因素被破坏,蛋白质的胶体性质就会被破坏,从而发生沉淀反应。

蛋白质的胶体性质具有重要的生理意义。在生物体中,蛋白质与大量水结合构成各种流动性不同的胶体系统。实际上,细胞的原生质就是一种复杂的胶体系统,体内的许多代谢反应就在此系统中进行。

4. 沉淀反应

凡是能破坏水化膜和/或能中和表面电荷的物质均会导致溶液中的蛋白质发生沉淀。导致蛋白质发生沉淀的因素有：既破坏水化膜又中和电荷的中性盐、中和电荷的等电点 pH、破坏水化膜的有机溶剂、中和电荷的生物碱等。不导致蛋白质变性的沉淀方法经常用于蛋白质的分离、纯化，主要包括以下几种：

（1）盐析　在蛋白质溶液中加入一定量的中性盐，可使蛋白质溶解度降低并沉淀析出的现象称为盐析（salting out）。发生盐析的原因是盐在水中解离后，与蛋白质争夺水分子，破坏了蛋白质颗粒表面的水化膜。另外，离子可大量中和蛋白质表面上的电荷，使蛋白质成为既不含水化膜又不带电荷的颗粒而聚集沉淀。盐析时所需的盐浓度称为盐析浓度，一般用饱和百分比表示。不同蛋白质的分子大小及带电状况各不相同，盐析所需的盐浓度也不尽相同。因此，可以通过调节盐浓度，使混合液中不同的蛋白质分别沉淀析出，从而达到分离的目的，这种方法称为分段盐析。硫酸铵是盐析中最常用的中性盐。

Quiz1 为什么盐析实验常用硫酸铵，而不是氯化钠或氯化钾？

有时，在蛋白质溶液中加入的中性盐的浓度较低时，蛋白质溶解度不降反增，这种现象称为盐溶（salting in）。产生盐溶现象的原因是蛋白质颗粒上吸附某种无机盐离子后，蛋白质颗粒带同种电荷而相互排斥，同时与水分子的作用得到加强，从而导致溶解度提高。

（2）p*I* 沉淀　当蛋白质溶液处于 p*I* 时，蛋白质分子主要以两性离子形式存在，净电荷为零。此时蛋白质分子失去同种电荷的排斥作用，很容易聚集而发生沉淀，这种沉淀蛋白质的方法叫 p*I* 沉淀。p*I* 沉淀既可用来分离蛋白质，也可以用来粗略测定某种蛋白质的 p*I*。

（3）有机溶剂引起的沉淀　某些与水互溶的有机溶剂（如甲醇、乙醇、丙酮等）可使蛋白质产生沉淀，是因为这些有机溶剂和水的亲和力大，能破坏蛋白质表面的水化膜，从而使蛋白质的溶解度降低。此法也可用于蛋白质的分离、纯化。

（4）重金属盐作用造成的沉淀　当蛋白质溶液的 pH 大于其 p*I* 时，蛋白质带负电荷，这时如果遇到重金属，会与重金属离子结合形成不溶性的蛋白盐而沉淀。由于重金属的作用通常会使蛋白质失活，因此很少用它来纯化蛋白质。

Quiz2 为什么重金属会导致蛋白质失活？

5. 蛋白质变性

蛋白质变性是指蛋白质受到某些理化因素的作用，其高级结构受到破坏（去折叠）、生物活性随之部分或全部丧失的现象。

蛋白质变性理论最早于 1931 年由北京协和医学院生化学家吴宪提出："天然蛋白质之分子，因环境种种之关系，从有程序而坚密之构造，变为无程序而散漫之构造，是为变性作用。"蛋白质的构象直接决定其功能，某些外界因素改变了蛋白质的独特构象，因而发生变性。蛋白质之所以容易发生变性，是因为维持蛋白质高级结构的作用力主要是次级键。当维系蛋白质高级结构的次级键受到严重破坏时，蛋白质必然发生变性。

蛋白质的变性与水解是不同的。蛋白质变性后一级结构没有发生任何变化，只是高级结构发生了变化，而水解则导致了肽键的断裂。

导致蛋白质变性的物理因素有加热、加压、机械作用（如搅拌、振荡）、表面张力、超声波和辐射等；化学因素包括强酸、强碱、高浓度盐、重金属盐、疏水分子、尿素和有机溶剂（如乙醇和氯仿）等。

疏水分子通过扰乱蛋白质分子内的疏水作用而导致变性。例如，长链脂肪酸的碳氢链不仅能与蛋白质疏水口袋非特异性结合，还能破坏蛋白质分子内部的疏水作用，从而抑制多种酶的活性。为了防止这种情况在细胞内发生，体内的长链脂肪酸和其他高度疏水性分子都有专门的结合蛋白与之结合。

蛋白质变性是一个复杂的过程，其中可能会出现一些不稳定的中间物。少数蛋白质的变性是可逆的，即在特定的条件下（变性因素解除以后）可以恢复到原来的构象，其生物活性也随之恢复，这就是蛋白质的复性。但大多数蛋白质的变性是不可逆的。例如，煮熟的鸡蛋即便冷却也绝对不会再变

回原来的生鸡蛋。

蛋白质变性以后，其理化性质也发生一系列的变化。这些变化可以作为检测蛋白质变性的指标。主要的变化包括以下几点：①溶解度降低。这是因为变性导致蛋白质内部的疏水基团外露。但变性蛋白质不一定都沉降，而沉降出的蛋白质也不一定变性。②黏度增加。这是因为蛋白质变性以后，肽链变得更加伸张，使长宽比提高。③生物活性丧失。例如，酶变性后丧失催化功能，这是蛋白质变性最重要的标志。④更容易被水解。这是因为多肽链构象变得更为松散和伸展，肽键更容易受到酸、碱或蛋白酶的作用。例如，胃酸除了能让微生物体内的蛋白质变性，从而杀死微生物以外，还能让食物中的蛋白质变性，以利于消化道内各种蛋白酶对蛋白质的消化。⑤结晶行为发生变化，通常是让蛋白质丧失结晶能力。

蛋白质的折叠状态与去折叠状态的能量差异比较小，因此有时仅仅一个点突变就能显著改变一种蛋白质对热的稳定性。蛋白质的温度敏感型突变体（temperature sensitive mutant）更容易发生热变性，而使用这样的突变体可以帮助鉴定某些蛋白质在细胞内的功能。

Quiz3 你如何使用温度敏感型突变体来确定一种蛋白质的功能？

某些蛋白质经突变以后热稳定性可能会提高。另外，从一些生存在极端环境中（如嗜热、嗜酸、嗜碱、嗜盐、嗜高压等）的微生物体内得到的蛋白质往往能够抵抗极端因素的作用。例如，从一种嗜热菌中提取出的 DNA 聚合酶能抵抗 100℃的高温，在 72℃时其活性最高，该酶现在已被广泛用于聚合酶链式反应（polymerase chain reaction，PCR）。这些蛋白质之所以能够抵抗极端因素作用，是因为在它们的高级结构中含有更多的次级键。

蛋白质变性在现实生活中具有重要意义。在临床上或工作中经常使用加热、酒精、紫外线等来消毒、杀菌，实际上就是使病毒和细菌的蛋白质变性而使其失去致病性和繁殖能力。在急救重金属盐中毒时也常常利用这一特性。例如，汞中毒时，早期可以服用大量富含蛋白质的乳制品或鸡蛋清，以使摄入的蛋白质在消化道与汞盐结合，形成变性的不溶物，从而阻止有毒的汞离子被消化道吸收进入人体，然后再通过洗胃等方法将不溶物洗出。

6. 蛋白质的水解

蛋白质在强酸、强碱或蛋白酶的催化下均能够发生水解。但需要注意的是，酸水解会破坏几种氨基酸，特别是 Trp 几乎全部被破坏，其次是 3 种羟基氨基酸（Ser，Thr，Tyr）。另外，Gln 和 Asn 在酸性条件下，容易水解成 Glu 和 Asp。酸水解常用硫酸或盐酸，使用最广泛的是盐酸；碱水解会导致多数氨基酸遭到不同程度的破坏，且产生消旋现象（racemization），但不会破坏 Trp；酶水解效率高，不产生消旋作用，也不破坏氨基酸，但不同的蛋白酶对肽键特异性不一样，因此，由一种酶水解获得的通常是蛋白质的部分水解产物。

Quiz4 一级结构为 WDNA 的四肽完全酸水解的产物是什么？

根据被水解肽键的位置，蛋白酶可以分为只能水解肽链内部肽键的内切蛋白酶和专门水解肽链末端肽键的外切蛋白酶。外切蛋白酶又可以分为专门水解 N 端肽键的氨肽酶（aminopeptidase）和专门水解 C 端肽键的羧肽酶（carboxypeptidase）。

7. 蛋白质的颜色反应

蛋白质分子中的肽键或某些氨基酸的 R 基团，可与某些试剂产生颜色反应，这些颜色反应经常被用来对蛋白质进行定性和 / 或定量分析。

目前常用的蛋白质定量法，如双缩脲法、福林 – 酚试剂法、考马氏亮蓝法等都是利用了蛋白质的颜色反应。此外，还有黄色反应、米伦氏反应、乙醛酸反应、坂口反应和醋酸铅反应等。其中，双缩脲法是基于蛋白质分子中的肽键，凡具有两个以上肽键的肽均能发生此反应，该方法受蛋白质特异氨基酸组成的影响较小，适用于毫克级蛋白质的测定；福林 – 酚试剂法即所谓的 Lowry 法，灵敏度高，但如果样品和标准蛋白质的芳香族氨基酸差异较大时，会有较大的系统误差；考马氏亮蓝 G–250 法即 Bradford 法，操作简单，灵敏度高。

e4–1 蛋白质的各种颜色反应

第二节 蛋白质一级结构的测定

蛋白质一级结构测定是研究蛋白质其他层次的结构和蛋白质功能的基础。测定蛋白质一级结构的方法有直接测定法和间接测定法。

一、间接测定法

间接测定法是先得到某一种蛋白质基因的核苷酸序列,然后根据通用的遗传密码表间接推导出由其决定的氨基酸序列。

如果是原核生物,可先直接从它的基因组 DNA(genomic DNA)中得到目标蛋白的基因,然后测定基因的碱基序列,找出可读框(open reading frame,ORF),最后根据遗传密码反推出氨基酸序列;如果是真核生物,可以先得到目标蛋白的 cDNA,然后测定 cDNA 的碱基序列,找出 ORF,最后同样根据遗传密码反推出氨基酸序列。真核生物一般难以直接从基因组 DNA 得到基因序列来推断其决定的氨基酸序列,因为真核生物的蛋白质基因内部大多含有不决定任何氨基酸序列的内含子(intron)(参看第三十七章"转录后加工")。

间接测定法的优点是快速,不需要纯化蛋白质,与直接测定多肽链的氨基酸序列相比,测定 DNA 的碱基序列要容易得多。但其缺点是,无法确定经后加工的蛋白质的最终序列,无法确定修饰的氨基酸,也得不到任何二硫键的信息。

间接测定法对含量低、不容易纯化的蛋白质很有用。许多难以纯化的膜内在蛋白都是用这种方法最先得到它们的一级结构的。另外,对于一些未知的蛋白质,间接测定法也很有用。例如,人们本来并不知道有神经珠蛋白的存在,更不知道它只在脑和视网膜中表达。但在分析脑 cDNA 库序列的时候,发现其中的一种 cDNA 似乎编码一种与 Mb 相似的蛋白质。在使用这种 cDNA 制作的探针对脑细胞进行检测以后,证明了脑细胞的确能表达神经珠蛋白的 mRNA。

二、直接测定法

直接测定法前后需要 9 大步,依次是:

1. 纯化目标蛋白

纯度应在 97% 以上(蛋白质纯化和纯度鉴定见本章第三节)。

2. 拆分肽链

如果目标蛋白含有 2 条或 2 条以上不同的肽链,必须先进行拆分,然后纯化出各条单链,再进入下一步,分别测定各条肽链的序列。

多亚基蛋白中亚基之间通常以次级键相连,因此在受到极端 pH、8mol/L 尿素、8mol/L 盐酸胍(guanidinium hydrochloride)或者高浓度盐的作用时,一般就能发生解聚。一旦解聚,可根据大小或电荷的差异将它们彼此分离。有时,亚基之间通过二硫键相连,这时需要利用下面的方法切开二硫键以后,再进行分离。

3. 打破二硫键

有许多方法可用来切开二硫键。Anfinsen 曾经使用巯基乙醇将胰核糖核酸酶的链内二硫键还原,使其破坏。这种方法也适用于亚基之间的链间二硫键。另外也可以用还原型的二硫苏糖醇(dithiothreitol,DTT)代替巯基乙醇来还原二硫键。在利用这两种还原剂来打破二硫键的同时,还需要使用碘代乙酸(iodoacetate)或 γ- 溴代丙胺(γ-bromopropylamine)将还原出来的巯基再氧化,以防止它们重新形成二硫键(图 4-1)。

另一种切开二硫键的方法是使用过甲酸(performic acid)直接氧化二硫键,将本来通过二硫键相连的 Cys 残基氧化成磺酸基 Cys。带负电荷的磺酸基之间的静电排斥可阻止二硫键重新形成。

(1)

$H-C(=O)-O-O-H$ 过甲酸

SO_3^- 磺酸基 Cys残基

(2)

$2HSCH_2CH_2OH$ β-巯基乙醇

SH SH

$S-CH_2-CH_2-OH$ | $S-CH_2-CH_2-OH$

(3)

ICH_2COOH 碘代乙酸 → HI +

$S-CH_2-COO^-$ S-羧甲基衍生物

(4)

$BrCH_2CH_2CH_2NH_2$ γ-溴丙胺 → HBr +

$S-CH_2CH_2CH_2-NH_2$

图 4-1　蛋白质二硫键的还原(1,2)和巯基的化学保护(3,4)

4. 分析各单链的氨基酸组成

测定蛋白质完全水解的产物可得到肽链的氨基酸组成。蛋白质的酸水解通常是将样品与 6 mol/L HCl 一起放在密封的玻璃小瓶中,在 110℃ 下放置 24、48 或 72 h。在酸水解条件下,Trp 几乎完全被破坏,因此它的含量需要用其他方法单独确定。尽管羟基氨基酸在酸水解条件下也会被破坏,但被破坏的速率较慢,因此可以将在 3 个时间段(24、48、72 h)得到的数据(3 种羟基氨基酸各自的含量)回推到零时段,从而还原出它们原来的含量。酸水解下释放出来的铵可用来估计蛋白质样品中 Asn 和 Gln 的总含量,但无法得到 Asn 和 Gln 各自的含量。于是,原来的 Asp 和 Glu 的含量也无法测出。因此,通常用 Asx 表示 Asp+Asn,Glx 表示 Glu+Gln。

反应结束以后,将反应混合物上柱分离。使用的分离方法可以是离子交换层析,也可以是反相 HPLC。如果使用离子交换层析,则将洗脱出来的氨基酸与茚三酮反应进行定量分析;如果是反相 HPLC,需要将氨基酸在上柱之前与 Edman 试剂反应,使其转变成 PTH- 氨基酸。无论是哪一种分离手段,现在都与定量反应一起完全实现了自动化,在专门的氨基酸分析仪中(amino acid analyzer)进行(图 4-2)。一个 3×10^4 的蛋白质通过自动分析仪只需要不到 1 h 的时间,需要的蛋白质样品量仅 6 μg(0.2 nmol)就足已。

氨基酸的组成分析并不能直接给出一个多肽的每一种氨基酸残基的数目,但能给出各种氨基酸

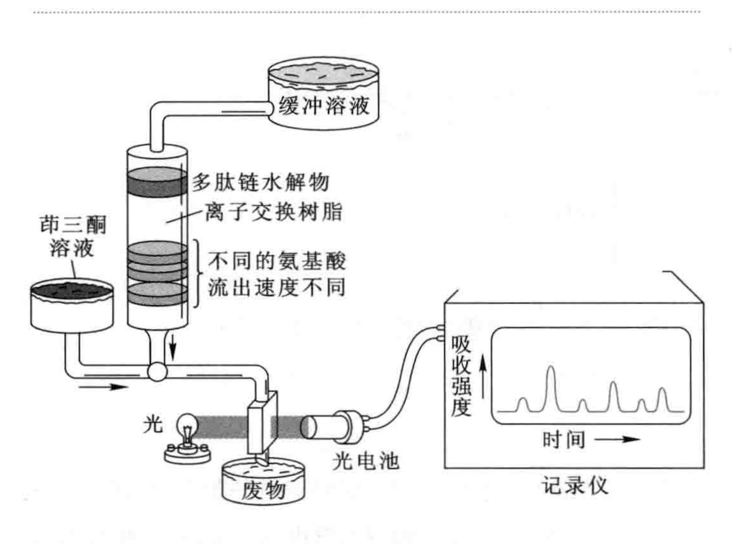


图 4-2 氨基酸分析仪的结构和原理

残基之间的相对比例。如果蛋白质的大小和样品中蛋白质的精确含量也能够确定，那么蛋白质分子中氨基酸的摩尔比就可以计算出来，在此基础上，就可以预测出特定蛋白酶的切点或 CNBr 切点的数目(参看第 6 步)。

5. 末端氨基酸残基的鉴定

先测定蛋白质末端的氨基酸，可以搞清楚目标蛋白有几个末端，进而有可能推断出它有几条肽链。如果一个蛋白质由 2 条或 2 条以上的肽链组成，那么测定出的 N 端或 C 端氨基酸很可能不止一种。但如果用常规的方法测定不出一种蛋白质的 N 端或 C 端氨基酸，那很可能是因为 N 端氨基或 C 端羧基被一些特殊的基团封闭住了。这时就需要用一些特殊的试剂来解除封闭。

关于 N 端氨基酸测定的方法在氨基酸化学反应中已有介绍(使用 Edman 试剂或 Sanger 试剂)，这里只描述 C 端氨基酸残基的测定。

C 端氨基酸的测定有两种方法：一种是酶法，另一种是化学法。酶法测定使用羧肽酶。有 4 种羧肽酶可供选择，即羧肽酶 A、B、C 和 Y，它们的特异性参看表 4-1。羧肽酶可以连续地从 C 端释放出氨基酸，直至遇到它不能水解的氨基酸。显然，越靠近 C 端的氨基酸残基释放的速率越快，利用这一点就可将 C 端氨基酸残基与其他氨基酸残基区分开来。目前已有人将羧肽酶 Y 进行改造，使之能够像 Edman 降解一样，从 C 端向 N 端连续测定肽段的氨基酸序列。

Quiz5 AKGRPA 经羧肽酶 A、B 和 C 水解后的产物分别是什么？

表 4-1 四种羧肽酶的来源和特异性

羧肽酶	来源	特异性
A	牛胰腺	C 端除 R、K 和 P 以外的氨基酸，但相邻的 N 端氨基酸不可以是 P
B	牛胰腺	C 端只能是 R 和 K，但相邻的 N 端氨基酸不可以是 P
C	柑橘叶	C 端可以是任何氨基酸，最适 pH 为 3.3
Y	酵母	C 端可以是任何氨基酸，但如果是 G，水解速率较慢

化学法又分为还原法和肼解法(hydrazinolysis)。还原法使用硼氢化钠将 C 端氨基酸残基转变成氨基醇，从而能够将它与其他氨基酸区别开来；肼解法使用无水肼(hydrazine)在 90℃下处理蛋白质近 20~100 h，这样的处理切开了蛋白质分子上所有的肽键，产生氨脂酰肼(amino-acyl hydrazide)，唯独 C 端氨基酸以游离的形式出现(图 4-3)。通过比较肼解产物与标准氨基酸的层析行为，可以确定出 C 端氨基酸(其层析行为不变)。由于高比例的肼能诱导副反应，这种方法一般只用于末端抵抗羧肽酶水解的肽。

6. 将肽链切成小的片段，再测定各小片段的氨基酸序列

Edman 降解能重复多次，且产率很高，但其应用仍然有限制：通常获得 20~30 个氨基酸序列没有问题，但要获得 50 个氨基酸以上的序列就很困难了！这是因为在每一轮循环中，PITC 不可能与 100% 的 N 端残基起反应，因而在多轮循环以后，错误会越来越多，那时就难以得到准确而有意义的序列信息了。

许多肽链的长度在 50 个氨基酸以上，因而在对它们进行 Edman 降解之前，需要将长的肽链进行特异性地切割，以产生长度低于 20~30 个氨基酸的肽段。

$$NH_2-CH(R_1)-C(=O)- \cdots HN-CH(R_{n-1})-C(=O)-N(H)-CH(R_n)-C(=O)-OH$$

$$\downarrow H_2N-NH_2$$

$$H_2N-CH(R_1)-C(=O)-NH-NH_2 + \cdots + HN-CH(R_{n-1})-C(=O)-NH-NH_2 + H_2N-CH(R_n)-C(=O)-OH$$

图 4-3 氨基酸的肼解反应

特异性切割多肽链的方法有酶法和化学法。酶法使用的是具有一定特异性的蛋白酶,这些蛋白酶对形成肽键的氨基酸残基的R基团有不同的要求,例如,胰蛋白酶只水解由Arg或Lys提供羧基的肽键。化学法则使用特异的化学试剂切开肽键,但有的试剂对肽键无选择性,如酸水解或碱水解,这样的试剂对肽链进行随机切割,用途不大。而能派上用场的一般是对肽键具有一定选择性的化学试剂。例如,溴化氰(CNBr)专门切割Met残基提供羧基的肽键。再如,羟胺(hydroxylamine)在pH 9时选择性切割Asn-Gly之间的肽键,而在温和的酸性条件下,只切割Asp-Pro之间的肽键。

e4-2 蛋白酶特异性的比较

无论是酶法还是化学法,切割后的产物都需要进行分离、纯化。分离后得到的每一种肽段还要单独进行氨基酸组成分析、末端氨基酸测定和全序列测定,以积累重建一个蛋白质全序列所必要的信息。

测定肽段序列的方法有Edman降解和质谱。关于Edman降解可参看第一章第二节“与PITC有关的反应”。质谱的基本原理是:待测样品在特定的条件下可转变为高速运动的离子。这些离子根据质荷比(质量与电荷之比)的不同在静电场和磁场的作用下得到分离,使用特定的检测器可记录各种离子的相对强度并形成质谱图。根据质谱图,不仅可以确定分子的结构、大小,还可测定蛋白质的氨基酸序列。

除了Edman降解和质谱以外,还有一种从肽段C端进行降解而获得序列的方法。这种方法是对Edman降解测序很有用的补充,特别是在遇到N端氨基被封闭而无法直接进行Edman降解的时候。此外,C端测序得出的信息还有助于对蛋白质翻译后C端剪切位点的鉴定、基因序列的验证和重组DNA技术中寡聚核苷酸探针的设计。

e4-3 C端降解

7. 选择不同的切点,重复步骤6

8. 根据片段重叠法,推断出肽链的全序列

在得到通过两种或两种以上切割方法产生的所有肽段的氨基酸序列以后,下一步要做的工作就是将各片段的序列放在一起进行比较,找出部分重叠的序列,在此基础上重建一种蛋白质完整的一级结构(图4-4)。

9. 二硫键的定位

如果一种蛋白质分子含有二硫键,则还需要对二硫键进行准确定位。对二硫键定位的基本步骤是:保留目标蛋白上的二硫键,直接用一种蛋白酶水解;找出含有二硫键的肽段以后,再用前面叙述的方法将二硫键拆开,分别测定两个肽段的顺序;在将它们的顺序与已测出的蛋白质一级结构进行比较以后,就能确定相应的二硫键位置。因此,“挑出”含有二硫键的肽段成为确定二硫键位置的关键。

寻找含有二硫键肽段的最佳方法是对角线电泳(diagonal electrophoresis)。按照图4-5所示进行操作,先将蛋白质部分水解物样品弥散地点在滤纸的一端,并开始进行第一向电泳。电泳结束以后,将样品条剪下,置于装有过甲酸的器皿中,用过甲酸蒸气处理2 h,使二硫键断裂,此时含有二硫键肽段的静电荷发生了改变(负电荷增加)。然后,将滤纸条附着于另一张新的滤纸上,进行第二向电泳,电泳条件与第一向完全相同,只是与第一次的方向成直角。在第二向电泳中,那些不含有二硫键的肽段电

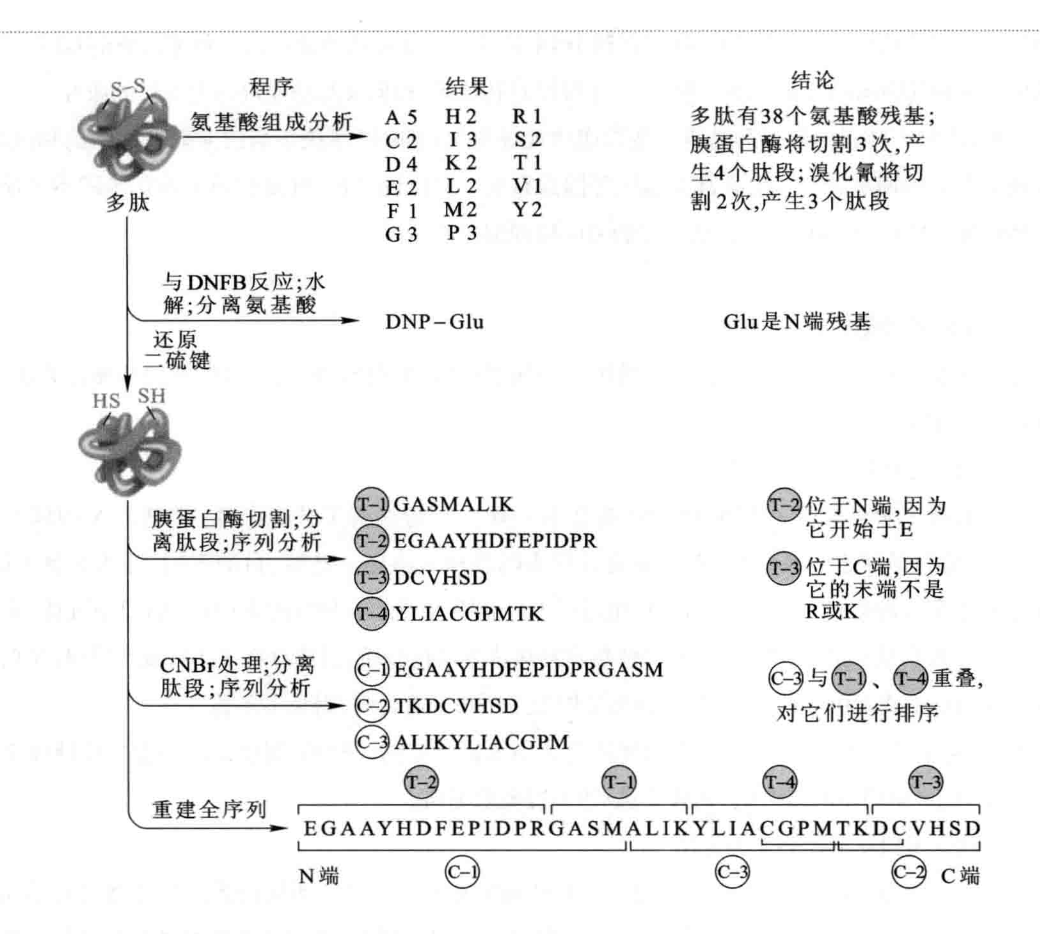


图 4-4 多肽序列分析实例（Nelson 等，2008）

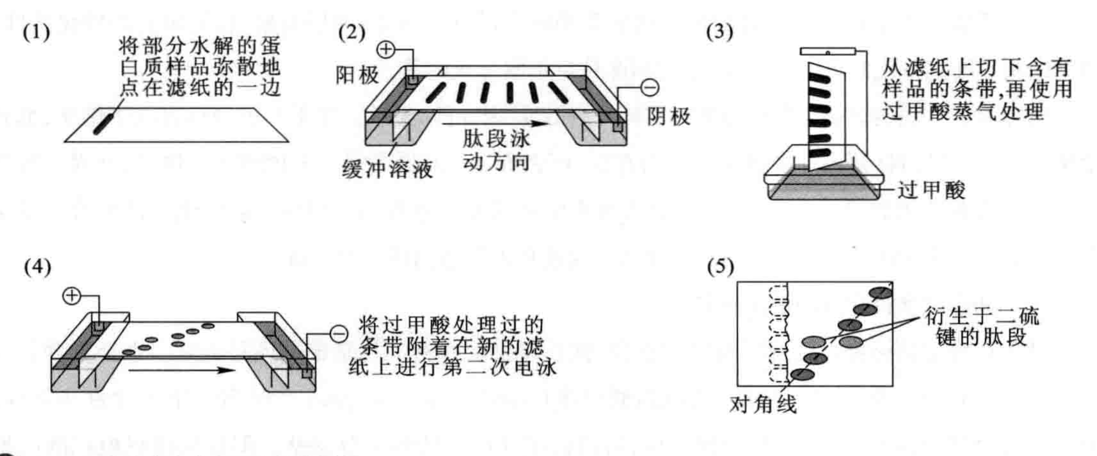


图 4-5 对角线电泳确定二硫键的原理

泳情况与第一向相同，电泳后均位于滤纸的对角线上；而那些含有二硫键的肽段由于电荷发生变化，其电泳速率就与第一向不同，从而导致这些肽斑偏离对角线。肽斑可用茚三酮显示。

第三节 蛋白质的分离、纯化和分析

蛋白质的分离和纯化不仅有助于研究蛋白质本身的结构和功能，还有助于研究其基因的结构与功能。同时，对于具有工业或药用价值的蛋白质也需要得到它们的纯品。例如，如果无法直接得到一种蛋白质的基因，那么通过纯化蛋白质，就可以测定出它的一部分氨基酸序列，然后根据遗传密码表设计和制作核酸探针，去“钓”出它的基因。或者以纯化的蛋白质免疫动物得到抗体，从 cDNA 表达文

库中分离出它的基因(参看第四十四章“重组 DNA 技术”)。如果已经分离出一种蛋白质的基因,只有在纯化到这种基因编码的蛋白质产物以后,才可以对其性质、构象及其功能进行全方位的研究。

然而,纯化一种蛋白质并非易事。能否成功地纯化蛋白质,既取决于蛋白质固有的结构和性质,也与纯化人员的创意、技巧、耐心甚至运气等因素有关。可以说,每一种蛋白质的纯化步骤不可能是完全相同的,但每一种蛋白质纯化所应用的方法原理都差不多。

一、蛋白质纯化的准备工作

在进行蛋白质纯化实验之前首先要解决三个问题:纯化蛋白质的目的;目标蛋白的测活方法;纯化蛋白质的原料。

(一) 明确纯化蛋白质的目的

不同的研究者纯化蛋白质的目的可能会不一样。有的是为了获得晶体,去进行 X 射线衍射分析;有的是为了测定一级结构;有的是将其作为药品投入市场。显然,目的不同,需求量就不同。如进行测序或质谱分析,约需要 1 pmol(相当于 50 ng 的 M_r 为 5×10^4)的蛋白质;如制备抗体,需要 200~500 μg 蛋白质;如做荧光分析、圆二色性和测热法等结构分析,需要量在 1~10 mg;如果做 X 射线衍射,需要 10mg 或者更多的蛋白质;而如果是作为药用和工业之用,则是多多益善。

此外,纯化蛋白质目的不同,对蛋白质纯度的要求也会不同。例如,测序使用的蛋白质纯度至少要在 97% 以上,而药用的蛋白质(如胰岛素)的纯度要求更高。

(二) 建立蛋白质活性测定的方法

在纯化一种蛋白质之前,首先需要建立一种检测它存在的方法。不同的蛋白质通常具有不同的活性,因此,检测它是否存在以及存在多少可以借助测定它特有的活性以及活性的高低来进行。换句话说,检测一个蛋白质的方法仅仅是以一种定量的手段测定一个特定活性存在的量。

蛋白质活性测定的方法通常根据它的生物功能来设计。例如,如果是酶,就测定它的催化活性;如果是刺激细胞分化的蛋白质,就测定它刺激特定细胞分化的能力。

选择测定方法的时候,要牢记两个标准:①测定是否方便、灵敏。如果一种测活方法不敏感,那就意味着需要大量目标蛋白才能显示它是否存在,而做到这一点通常是十分困难的。同样,如果一种测活方法需要花几天的时间才能完成,就会大大影响到实验的进程,并且时间越长,蛋白质越容易发生水解。②测定的方法是否高度专一。需要防止功能相近的蛋白质产生干扰。

(三) 选择富含目标蛋白的原材料

对含有特定目标蛋白的起始材料的选择,往往直接关系到纯化是成功还是失败。选择富含目标蛋白的原材料是一般原则。例如,乙酰胆碱受体(acetylcholine receptor)在肌肉细胞中含量非常少,因此,如果想从肌肉中分离乙酰胆碱受体,可以说成功的可能性十分渺茫。但如果选择电鳐的电板(electroplax)作为原材料,纯化就要容易得多,这是因为电板细胞的一面密集了大量的乙酰胆碱受体。再如,从肾细胞中纯化水孔蛋白(aquaporin)是一种很好的选择,因为肾与水代谢有密切关系,构成其细胞的质膜上应该富含这种蛋白质。

有时,某些蛋白质的含量本来就很低,难以找到富含它的组织,这时可考虑先得到它的基因,然后利用基因工程进行高表达后再进行分离、纯化。

二、蛋白质纯化的一般注意事项

在进行任何一种蛋白质纯化的时候,都要时刻注意维护它的稳定,保护它的活性,有一些通用的注意事项需要牢记,它们包括以下几点:①操作尽可能在低温条件下(如置于冰上或者在冷库内)进行。②待纯化材料不要太稀,蛋白浓度维持在 μg/mL~ mg/mL。③合适的 pH,除非是进行聚焦层析或等电点沉淀,所使用的缓冲溶液的 pH 避免与 p*I* 相同。④使用蛋白酶抑制剂,防止蛋白酶对目标蛋白

的降解。⑤避免样品反复冻融和剧烈搅动，以防蛋白质的变性。⑥缓冲溶液成分尽量模拟细胞内环境。⑦在缓冲溶液中加入 0.1~1 mmol/L DTT（或 β- 巯基乙醇），防止蛋白质的氧化。⑧加入 1~10 mmol/L EDTA 金属螯合剂，防止重金属对目标蛋白的破坏。⑨使用灭菌溶液，防止微生物生长。

三、蛋白质纯化的常见方法

分离纯化蛋白质的各种方法都是利用不同蛋白质在理化性质上的差别，例如溶解性、相对分子质量和形状、表面电荷、表面疏水性、与特定配体结合的性质等方面。

常用的方法有：沉淀（利用蛋白质的溶解性）；离子交换（利用蛋白质的电荷性质）；聚焦层析（利用蛋白质的电荷性质）；凝胶过滤（利用蛋白质的相对分子质量和形状）；疏水作用层析（利用蛋白质的疏水性）；亲和层析（利用蛋白质与特定配体的特异性结合性质）。

在纯化过程中，有时候样品体积较大，需要先进行浓缩以缩小体积。常见的浓缩方法有：沉淀、冻干、反透析、超滤（ultrafiltration）。

有关层析特别是离子交换层析已在第一章中作了介绍，这里主要就其他几种方法的原理作简单介绍，有关详细的操作过程可以参考有关实验手册。

（一）沉淀与离心

沉淀是根据不同蛋白质在特定条件下溶解性不同，而对它们进行选择性沉降从而达到分离目的的一种粗纯化方法。它通常用于将目的蛋白从大体积的粗抽取物中游离出来。这种方法既能除去许多杂质，又有浓缩之效。

实现选择性沉淀的方法包括：改变 pH 或改变离子强度（盐析）。无论是哪一种沉淀方法，产生的沉淀物都需要借助离心的手段才能与上清液分开。

离心方法是根据分子的特征密度来分离大分子的。如果一种颗粒的密度大于其介质溶液的密度，那么这种颗粒有可能通过溶液发生沉降。颗粒沉降的速率与颗粒和介质溶液之间的密度之差成正比。任何一种颗粒在离心力作用下通过溶液发生沉降的趋势可用沉降系数（sedimentation coefficient），即 S 表示：$S=(p-m)V/f$。这里 p 为颗粒或大分子的密度，m 为介质或溶液的密度，V 为颗粒体积，f 为摩擦系数。非球形分子具有更大的摩擦系数，因此其沉降系数较小。颗粒越小或形状与球形相差越远，在离心机内就沉降得越慢。

沉降系数对于生物大分子来说，多数在 $(1\sim500)\times10^{-13}$ s。为应用方便起见，人为规定 1×10^{-13} s 为一个 S 单位。一般单纯的蛋白质在 1~20 S，较大核酸分子在 4~100 S，更大的亚细胞结构在 30~500 S。

离心既可以作为分离和纯化大分子或亚细胞组分的制备技术，也可用来分析大分子的流体动力学性质。分析用超速离心机仅能对少量样品（少于 1 ml）进行分析；制备用离心机是专为提取、纯化大分子组分设计的，可对大量的样品（10~2 000 ml）进行分离。

制备用离心机的应用非常广泛，根据其性能可分为低速离心机（最大速率不超过 10 000 r/min）、高速离心机（最大转速为 20 000~25 000 r/min）和超速离心机（最大转速可超过 75 000 r/min）。高速离心机一般均配有冷却装置，用以防止转轴的温度过高而使生物样品失活。此外，所有超速离心机都装配真空装置，以减少离心室内转头与空气之间的摩擦作用。

制备用离心机在应用方法方面又有沉降离心和梯度离心两种形式：

（1）沉降速率离心（velocity sedimentation centrifugation） 在离心管内溶液密度均一的情况下进行的离心。最常见的方法是差速离心（differential centrifugation），这是对含两种以上大小不同的待分离物质的混合液，以不同离心速率分段离心沉淀，使之相互分离的离心方法。

（2）梯度离心 将待分离的物质在具有密度梯度的介质中进行的离心。梯度离心又可分为密度梯度离心和平衡密度梯度离心。

① 密度梯度离心 密度梯度离心（density gradient centrifugation）是在密度梯度介质中进行的一种

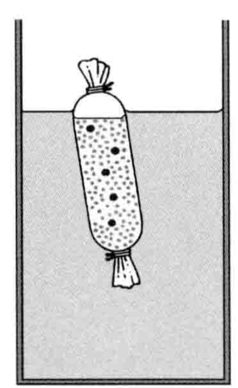
沉降速率离心,被离心的物质根据其沉降系数不同而进行分离。常用蔗糖或甘油来制备梯度。梯度的作用是稳定离心液以减少扩散,而得到较为锋利的区带。密度梯度离心和差速离心一样是依赖时间的,如果离心时间过长,已经分离的物质都可沉降到离心管的底部或某一密度区域。

② 平衡密度梯度离心　平衡密度梯度离心(equilibrium density gradient centrifugation)虽然也是在密度梯度介质中进行的,但被分离的物质是依靠它们的密度不同而进行分离的,此种离心常用 CsCl 等无机盐类制备密度梯度。在梯度介质中,当被分离的物质分别达到与其密度相同的介质部位时,就不再移动,从而达到分离的目的。平衡密度梯度离心是不依赖于时间的,只要梯度不破坏,时间延长也不影响分离。

Quiz6 DNA、RNA 和蛋白质的混合物经 CsCl 密度梯度离心分离以后,它们各自在离心管中的位置如何?

(二) 透析和超滤

透析(dialysis)是利用蛋白质等生物大分子不能透过半透膜(semipermeabl membrane),但小分子物质和离子能够通过半透膜而进行纯化的一种方法(图 4-6)。这种方法经常被用于去除大分子溶液中的小分子物质或改变蛋白质溶液的组成。


开始透析

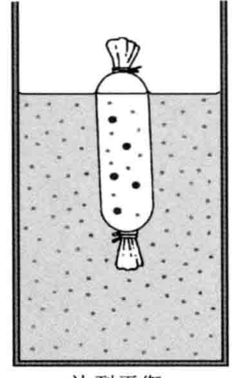

达到平衡

图 4-6　透析示意图

如果在装有蛋白质或其他大分子溶液的透析袋外,放入高浓度吸水性强的多聚物(如固体聚乙二醇),透析袋内的水便迅速被袋外多聚物所吸收,从而达到了浓缩袋内液体的目的。这种方法称为“反透析”。

超滤(ultrafiltration)对透析原理进行了改进,它利用具有一定大小孔径的微孔滤膜,在常压、加压或减压条件下对生物大分子溶液进行过滤,使大分子保留在超滤膜上面的溶液中,小分子物质及水过滤出去,从而达到脱盐、浓缩或更换缓冲液的目的。

(三) 电泳

电泳的原理已在氨基酸的分离技术中作过介绍,这里仅介绍 3 种与蛋白质分离纯化有关的电泳技术,即等电聚焦电泳(isoelectric focusing electrophoresis,IFE)、十二烷基硫酸钠聚丙烯酰胺凝胶电泳(sodium dodecyl sulfate polyacrylamide gel electrophoresis,SDS-PAGE)和双向电泳(two-dimensional gel electrophoresis)。

1. IFE

IFE 需要在凝胶中加入两性电解质,以便在两极之间建立 pH 梯度,处在其中的蛋白质分子在电场的作用下发生迁移,最后各自移动到并聚焦于凝胶上某一特定的位置,各个位置的 pH 与聚焦在此的蛋白质的 p*I* 相同。通过 IFE,不仅可以实现 p*I* 不同的蛋白质之间的分离,还可以直接测定出各种蛋白质的 p*I*。

2. SDS-PAGE

蛋白质在非变性聚丙烯酰胺凝胶(native PAGE)中电泳时,其迁移率取决于分子的大小、形状及其所带净电荷等因素。如果加入一种试剂使电荷因素消除,并使分子形状趋于一致,那电泳迁移率将只取决于分子的大小。在此基础上,还可以测定出蛋白质的相对分子质量。阴离子去污剂 SDS 具有这种作用。在向蛋白质溶液中加入足够量的 SDS 以后,SDS 即与蛋白质形成复合物。SDS 带负电,能使各种蛋白质 - SDS 复合物都带上相同密度的负电荷,其总量大大超过了蛋白质分子原来的电荷量,因而掩盖了不同种蛋白质间原有的电荷差别。SDS 还能打破维持蛋白质三级结构的次级键,使各种蛋白质都呈线状展开(图 4-7)。这样的蛋白质 - SDS 复合物,在凝胶电泳中的迁移率,不再受蛋白质原来的电荷和形状的影响,仅取决于它们的大小,因而 SDS-PAGE 可以用于测定蛋白质的相对分子质量。

SDS-PAGE 不仅可以用来测定蛋白质的大小,还可以用来鉴定蛋白质的纯度、定量蛋白质和二硫键的确定等。

3. 双向电泳

双向电泳也称二维电泳或 2D 电泳,它由第一向等电聚焦电泳和第二向 SDS - 聚丙烯酰胺凝胶电泳组成(图 4-8)。第一向利用等电聚焦的方法使 p*I* 不同的蛋白得到分离。将第一向电泳的凝胶转移

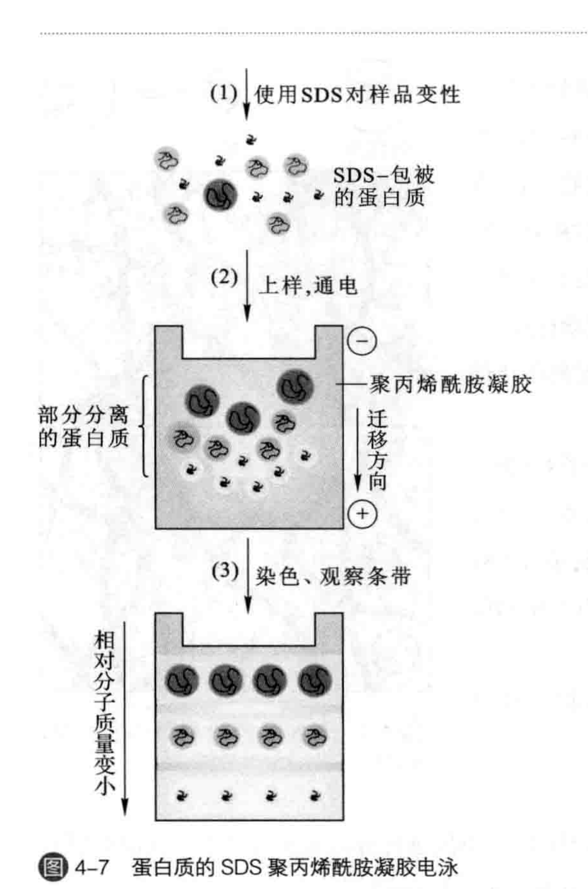


图 4-7 蛋白质的 SDS 聚丙烯酰胺凝胶电泳

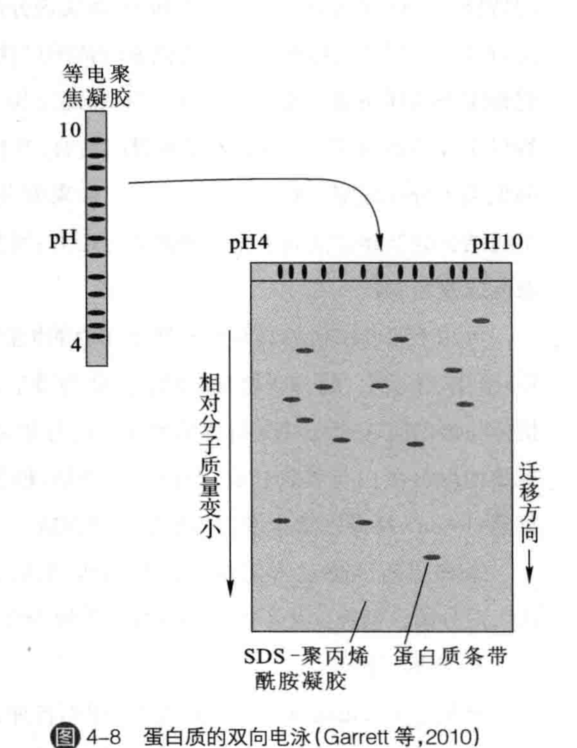


图 4-8 蛋白质的双向电泳(Garrett 等,2010)

到 SDS-PAGE 平板的顶部,即进行第二向电泳,这一次是按照分子大小来分离,两向结合便得到高分辨率的蛋白图谱。在对凝胶进行染色以后,可以观察到各个蛋白质的位置。

Quiz7 双向电泳的两向次序可以颠倒吗？为什么？

使用双向电泳,研究者有可能观测并建立一种特定细胞所表达的全部蛋白质的目录。ExPASy (http://expasy.hcuge.ch/)已建立了双向电泳数据库,数据库内含有多种类型的细胞和组织的蛋白质双向电泳图谱。

(四) 层析

层析在蛋白质分离纯化中是不可替代的一种方法,其基本原理参看第一章氨基酸。常见的层析方法及其比较见表 4-2。

表 4-2 几种层析分离方法的比较

方法	容量	纯化位置	成本	方法	容量	纯化位置	成本
离子交换层析	高	早期	低	凝胶过滤	底	后期	低
疏水层析	高	早期	中	亲和层析	高	任何时段	中
聚焦层析	中等	后期	高				

1. 疏水作用层析

疏水作用层析(hydrophobic interaction chromatography,HIC)是利用不同蛋白质在疏水性质上的差别,而对特定蛋白质进行纯化的一项层析技术。其固定相是连接到惰性基质上的疏水基团,如辛烷基(octyl)或苯环。其原理类似于盐析,即在高离子强度下使样品上柱(如 1 mol/L 硫酸铵),以此除去蛋白质的水化层,由于表面上亲水基团最少的蛋白质最容易失去水,也就最容易暴露出它们的疏水基团,并与树脂上的疏水基团结合;反之,表面上亲水基团最多的蛋白质最难失去水化层,也就最难与固定相结合。

HIC 的洗脱方法有:降低洗脱液的盐浓度、增加洗脱液的去污剂浓度或者改变 pH。

2. 亲和层析

亲和层析（affinity chromatography）是利用待分离物质和它的特异性配体间具有特异性的亲和力，而实现分离的一类特殊层析技术。其固定相是带有共价偶联配体的惰性树脂，当含有混合组分的样品通过此固定相时，只有和固定相上的配体具有特异亲和力的物质，才能被固定相吸附结合，其他没有亲和力的无关组分就随流动相流出（图 4–9）。如果改变流动相成分，可将结合的亲和物洗脱下来。洗脱的方法有：增加游离配体的浓度或改变 pH。

蛋白质　配体

图 4–9　亲和层析的原理（Garrett 等，2010）

可用于亲和层析的具有特异性亲和力的生物分子对（蛋白质－配体）主要有：酶与底物/抑制剂；受体与激素；抗体与抗原；抗体与蛋白质 A；带有 His 标签的融合蛋白与金属镍；带有 GST 标签的融合蛋白与谷胱甘肽；蛋白质与染料；糖蛋白与植物凝集素（lectin）；特异性结合蛋白与被结合的物质。

Quiz8 你如何设计亲和层析实验，来分离分子伴侣和乙醇脱氢酶?

亲和层析纯化过程简单、迅速、高效，有时能够"一柱到位"，对分离含量极少又不稳定的活性物质较为有效。

3. 聚焦层析

聚焦层析（chromatofocusing，CF）是根据各种蛋白质的 p*I* 不同进行分离的层析技术。其流动相为多元缓冲剂（polybuffer），固定相为多元缓冲交换剂（polybuffer exchanger）。

CF 的原理与等电聚焦和离子交换相似，不过是在含有 pH 梯度的离子柱内进行。当将特定的多元缓冲液自上而下流过填充在层析柱内的多元缓冲交换剂时，由于交换剂带有具缓冲能力的电荷基团，会在多元缓冲交换剂的表面从上到下自动地形成连续的 pH 梯度。因为蛋白质所带电荷取决于它的 p*I* 和层析柱中的 pH 值，所以如果使用的是阴离子交换剂，则当柱中的 pH 低于蛋白质的 p*I* 时，蛋白质带正电荷，而不与阴离子交换剂结合。但随着洗脱液向前移动，固定相中的 pH 随着洗脱时间延长而变化。当蛋白质移动到环境 pH 高于其 p*I* 时，蛋白质由带正电荷变为带负电荷，并与阴离子交换剂结合。当洗脱液流过的时候，蛋白质周围环境的 pH 再次低于 p*I*，于是它又带上正电荷，并从交换剂解离下来。随着洗脱液向柱底部的迁移，上述过程将反复进行，各种蛋白质就在各自的 p*I* 被洗下来，从而实现了分离的目的。

4. 凝胶过滤层析

凝胶过滤层析（gel filtration chromatography）又称为分子筛过滤层析（molecular sieve chromatography）或大小排阻层析（size exclusion chromatography）等。它是一种按相对分子质量来分离物质（分子形状也起一定作用）的层析方法。其固定相是装在层析柱内细而多孔的凝胶颗粒，例如葡聚糖（商品名为 Sephadex）、琼脂糖（商品名为 Sepharose）或聚丙烯酰胺（商品名为 Sephacryl）。在将样品加到充满着凝胶颗粒的层析柱以后，使用缓冲液洗脱。大分子无法进入凝胶颗粒的内部，只能存在于凝胶颗粒之间的流动相，因而以较短的路径和较快的速率首先流出层析柱，而小分子则能自由出入凝胶颗粒内部，并很快在两相之间形成动态平衡，于是要走更长的路径和花费更长的时间流过柱床。不同大小的分子因此得以分离（图 4–10）。

Quiz9 预测以下五种蛋白质通过凝胶过滤层析的次序：蛋白质 A（3×10^4）、蛋白质 B（5×10^4）、蛋白质 C（4×10^4）、蛋白质 D（7×10^4）和蛋白质 E（6×10^4）。

5. HPLC

HPLC 的层析原理与经典的液相层析及气相层析一致，只不过对固定相颗粒的大小、流动相的流速进行了很大的改进，使颗粒更细（3~10 μm）、流速更快，并实现了自动化。以反相 HPLC（reversed-phase HPLC）为例，其固定相由共价结合到硅上的碳氢链构成。它的纯化原理与疏水层析一样，蛋白质通过疏水作用结合到柱上，可用降低离子强度或增加洗脱液的疏水性的方法进行洗脱。

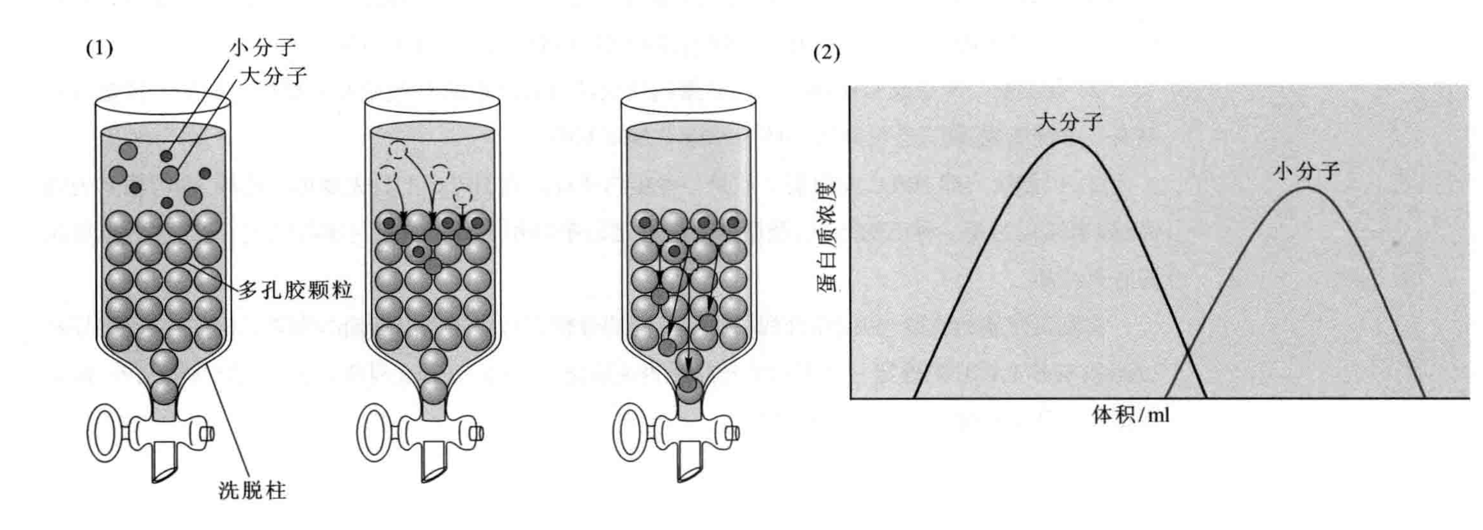


图 4-10 凝胶过滤的原理(Garrett 等,2010)

四、蛋白质纯化方案的设计

在纯化蛋白质之前,首先需要熟悉各种蛋白质纯化方法的原理,以及对纯化过程中的一些注意事项有所了解;然后就是选择好测活的方法和含有目标蛋白的原材料;接着就是设计纯化方案了。

为了纯化一种蛋白质,必须从起始物质开始就选择一种物理的或化学的途径进行分级分离。分级分离的方法有离心、盐析、离子交换层析、速率沉降、平衡密度沉降、亲和层析、疏水层析、聚焦层析和凝胶过滤等。先用哪一种后用哪一种没有固定的模式,一般靠的是经验与摸索。无论哪一种分离方法完成以后,都会发现某些部分有活性,某些部分无活性。这时需要做的就是保留和合并有活性的部分。分级的目的一是除去污染物,二是合并有活性的部分,来富集目标蛋白。被富集的部分再经历另外一轮分级处理,这样的过程还可以继续重复。如果蛋白活性能够被精确地测定,那每一次分级后得到的样品总量就可以确定,每一次制备的比活性(specific activity)即每毫克蛋白内的蛋白活性也可以计算出来。

在一个精心设计的纯化程序中,每一级分离都应该导致杂质的去除和比活性的提高。但在实际操作中,在比活性提高的同时,几乎总会有总活性(total activity)的损失。最大地提高比活性同时尽量减少总活性的损失是任何一种纯化程序的目标,而设计一种好的纯化方案通常是考虑到多种因素后的折中。一个纯化流程可以总结在一张纯化表(purification table)上(参看第十三章“酶的应用及研究方法”),它必须真实地记录各项有效的数据。这样的纯化表不仅是将来论文中有力的论据,也为评估一步纯化步骤是否合理提供了依据。

一个典型的蛋白质纯化方案开始于完整的组织,一般经过四个阶段:①破碎细胞或组织(混合和匀浆);②去除残渣(离心);③沉淀 / 浓缩(硫酸铵或聚乙二醇);④纯化(层析);⑤鉴定,包括纯度、大小或 p*I* 的测定。

在纯化过程中,使用哪些层析方法以及被选用的层析方法的前后次序可参看表 4-2。

五、蛋白质纯度的测定

纯净的蛋白质样品应该不含有其他杂蛋白和杂质。一般认为,当一种蛋白质被纯化到恒定的比活性或达到均一性(homogeneity)以后,就可被认为是纯品了。然而,单凭恒定的比活性是不够的,还需要使用其他的方法加以验证。

其他鉴定蛋白质纯度的方法有:

(1) 电泳法　如聚丙烯酰胺凝胶电泳、等电聚焦和毛细管电泳等。纯净的蛋白质电泳的结果应该是一条带。如果使用 SDS-聚丙烯酰胺凝胶电泳,就要特别小心,原因是 SDS 能够破坏蛋白质的四级

结构，而导致异源寡聚体蛋白质可能出现几条带。此外，电泳法检测蛋白质的纯度时，应取分布在蛋白质等电点两侧的两个不同的 pH 分别进行检测，这样得出的结论才更可靠。

(2) 化学法　N 端或 C 端测定。纯品蛋白质应该具有恒定的 N 端或 C 端组成。如果一种蛋白质只有一条链组成，就只会检测到一种 N 端或 C 端氨基酸。

(3) 仪器法　即 HPLC 或质谱。如果一种蛋白质样品在 HPLC 上只表现单一的峰，则可视其为纯品；如果纯化的是一种已知蛋白，经质谱分析测定出来的相对分子质量与实际的值一致，那么也可认为它是纯品。

检测蛋白质的纯度，必须综合两种原理不同的分析方法才能作出准确的判断。例如，仅仅用凝胶过滤和 SDS-PAGE 来确定一个蛋白质样品是否纯净是欠妥的，因为这两种方法的原理是相同的，都是根据相对分子质量的不同分离蛋白质。

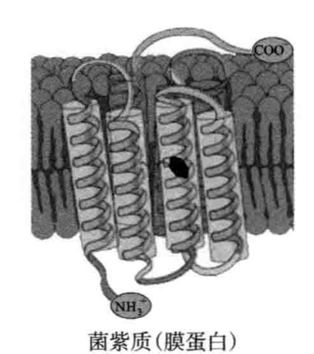
第四节　蛋白质的分类



胶原蛋白（纤维状蛋白质）



肌红蛋白（球状蛋白质）



菌紫质（膜蛋白）

图 4-11　根据分子形状、溶解性质以及是否与膜结合对蛋白质进行的分类

常见的蛋白质分类方法至少有五种，其分类的依据分别是溶解性、组成、形状、功能和三维结构。

按照溶解性质，蛋白质可分为七大类：①白蛋白（albumin）。又称为清蛋白，较小，能溶于水和盐溶液，可被饱和硫酸铵沉淀。如血液中的血清清蛋白和鸡蛋中的卵清蛋白等都属于清蛋白。②球蛋白（globulin）。通常不溶于水而溶于稀盐、稀酸或稀碱溶液，能被半饱和的硫酸铵沉淀。例如，大豆种子中的豆球蛋白（legumin）、血液中的血清球蛋白、肌肉中的肌球蛋白（myosin）以及免疫球蛋白都属于这一类。③组蛋白（histone）。可溶于水或稀酸，是染色体的结构蛋白，含有丰富的 Arg 和 Lys 残基，因此是一类碱性蛋白质。④精蛋白（protamine）。易溶于水或稀酸，是一类结构简单的蛋白质。与组蛋白一样，精蛋白也是一类碱性蛋白质，含有较多的碱性氨基酸，但缺少 Trp 和 Tyr。精蛋白存在于成熟的精细胞中，与细胞核 DNA 结合在一起，如鱼精蛋白。⑤醇溶蛋白（prolamine）。溶于 70%~80% 的乙醇，但不溶于水、无水乙醇或盐溶液，多存在于禾本科作物的种子中，如玉米醇溶蛋白、小麦醇溶蛋白。⑥谷蛋白类（glutelin）。不溶于水或稀盐溶液，但溶于稀酸或稀碱。谷蛋白存在于植物种子中，如水稻种子中的稻谷蛋白和小麦种子中的麦谷蛋白等。⑦硬蛋白类（scleroprotein）。不溶于水、盐溶液、稀酸、稀碱，主要存在于皮肤、毛发、指甲、蹄和蜘蛛丝等中，起支持和保护作用，如角蛋白（keratin）、胶原蛋白（collagen）、弹性蛋白（elastin）和丝蛋白等。

按照化学组成，蛋白质又可分为简单蛋白质(simple protein)和结合蛋白质(conjugated protein)两类。简单蛋白质也叫单纯蛋白质，它们仅由氨基酸组成，水解产物仅有氨基酸，不含其他化学成分。例如，胰核糖核酸酶。结合蛋白质也叫缀合蛋白质，它们除含有氨基酸以外，还含有非氨基酸成分作为其结构的一部分。非氨基酸成分可分为四种形式：第一种是氨基酸残基侧链上被修饰的各种基团，如糖基、脂酰基、甲基和磷酸根等；第二种是指与蛋白质结合的无机成分，通常是一些金属离子；第三种为辅酶，专指与蛋白质结合并不紧密的有机分子，如辅酶 I 和辅酶 II；第四种为辅基，专指与蛋白质紧密结合的有机分子，如 FAD 和 FMN。

根据非蛋白成分的性质，结合蛋白质还可以分为糖蛋白（glycoprotein）、脂蛋白（lipoprotein）、核蛋白（nucleoprotein）、色蛋白（chromoprotein）、金属蛋白（metalloprotein）、磷蛋白（phosphoprotein）、血红素蛋白和黄素蛋白等。这些结合蛋白的非蛋白质成分分别是共价结合的糖基、非共价结合的脂、非共价结合的核酸、共价结合或非共价结合的生色基团（如血红素）、配位结合的金属离子、共价结合的磷酸根、共价或非共价结合的血红素辅基以及共价或非共价结合的黄素（flavin）的衍生物。

按照分子形状、溶解性质以及是否与膜结合，蛋白质又被分为球状蛋白质（globular protein）、纤维状蛋白质（fibrous protein）和膜蛋白（membrane protein）三大类（图 4-11）。纤维状蛋白质结构伸展，呈

框 4-1　身边的生物化学——毒也美丽

若是肉或鱼烹调不当或腌制不当，一种致命的食物中毒——肉毒杆菌中毒可能会发生。其罪魁祸首是由肉毒杆菌分泌的肉毒杆菌毒素（botulinum toxin，BTX），它能阻止乙酰胆碱从突触前囊泡释放，致使没有神经递质到达肌肉细胞表面与受体结合，于是肌肉不再收缩。如果不及时治疗的话，患者随时可能死亡。肉毒杆菌毒素是由肉毒杆菌在繁殖过程中所产生的一种神经毒素蛋白，是毒性最强的天然物质之一。1 mg 肉毒毒素能杀死 2 亿只小鼠。

令人惊讶的是，经纯化的肉毒杆菌毒素具有宝贵的医用价值。医学上，它被用于治疗面部不自主的肌肉痉挛，例如因乙酰胆碱的释放失控引起的面部抽动。1986 年，加拿大的 Carruthers 夫妇发现，这种能麻痹肌肉神经的药物可以使过度收缩的小肌肉放松，而达到舒展皱纹的效果，对表情纹特别有效。

美国食品和药物管理局（FDA）早已批准化妆品使用 BOTOX（肉毒杆菌毒素），这意味着制造商现在可以大张旗鼓地宣传其让人青春再现的奇效。肉毒杆菌毒素已被用于除皱多年，目前欧美比较时髦的除皱技术就是局部注射肉毒杆菌毒素。这种毒素如果能在局部加以控制地使用，可阻止面部肌肉不受控制的收缩，从而减轻面部变形。目前它的使用正在迅速蔓延，特别是在演艺圈。但若是使用不当，会造成头痛、过敏、复视、表情不自然的不良反应。特别是如果使用过量，它有可能扩散到其他的地方而产生强烈的副作用，严重可致死。事实上，一些电影导演经常抱怨道，有些演员由于使用过多的肉毒杆菌毒素已难以显示出各种自然的表情。

纤维状，长 / 宽 >10，不溶于水或稀盐溶液，机械强度高，在生物体充当结构的角色，如胶原蛋白和角蛋白。球状蛋白质结构紧密，呈球状，长 / 宽≤3~4，溶于水，如肌红蛋白和胰蛋白酶等。膜蛋白则以一定的方式定位于各种生物膜上，可分为膜外在蛋白、膜内在蛋白和脂锚定蛋白（参看第十六章“脂质与生物膜”）。

根据结构和进化上的亲缘关系，又可将蛋白质分为家族、超家族和栏。①家族：在进化上具有明确的亲缘关系。一般说来，属于同一家族的蛋白质至少有 30% 氨基酸序列是相同的。②超家族：在进化上可能具有相同的起源。属于同一超家族的蛋白质在氨基酸序列上的差别可能较大。③栏：具有相同的二级结构、相同的排列和相同的拓扑学连接。属于同一栏的蛋白质并不需要具有相同的一级结构，但通常具有相同的生物学功能。

根据功能，蛋白质又可以分为酶、调节蛋白、运输蛋白、贮存蛋白、收缩和运动蛋白、结构蛋白、接头蛋白、保护和防御蛋白、毒蛋白和奇异蛋白等（参看第三章“蛋白质的功能及其与结构之间的关系”）。

第五节　多肽的固相合成

多肽的固相合成由美国科学家 Bruce Merrifield 于 1963 年创立。其基本原理是：根据事先设计好的多肽序列，先将所要合成肽链的 C 端氨基酸的羧基以共价键固定到一个不溶性的高分子树脂上，然后从 C 端向 N 端，重复“去保护（除去要连接的氨基酸 α- 氨基上的保护基团）→活化（激活前一个已和固相柱相连的氨基酸的羧基）→偶联（让去保护的氨基和活化的羧基之间缩合，形成肽键）→洗脱和过滤（去除没有反应完的各种试剂）”这几步反应，直至达到所要合成的肽链长度。然后再除去侧链上的保护基团，最后将肽链从树脂上裂解下来，经过纯化等处理，即可得到所要的多肽。

固相法最大的进步是主要解决了分离、纯化方面的难题。每步反应后只需简单地洗涤树脂，便可达到纯化目的，这克服了经典液相合成法中的每一步产物都需纯化的麻烦，为自动化合成肽链奠定了基础。Merrifield 也因开发多肽固相合成法获得 1984 年诺贝尔化学奖。

要知道，Merrifield 是在 1959 年的某一天，于 Rockefeller 大学的 Flexner 厅的一个电梯里，向他的一个同事提出他的想法的！至今那个电梯里还手写着一句话：SOLID-PHASE PEPTIDE SYNTHESIS WAS BORN HERE。

e4-4 多肽固相合成

科学故事——NGF 的发现

神经生长因子（nerve growth factor，NGF）是一种多肽生长因子，其作用是促进神经元的存活（neural survival）和分化。从 20 世纪 50 年代初开始，意大利神经生物学家 Rita Levi-Montalcini 一直在美国华盛顿大学路易斯分校的 Victor Hamburger 实验室工作，主要研究移植或去除鸡趾对神经系统发育的影响。她在研究中发现，移植的鸡趾不但能发育，而且还受到神经的支配（innervated）。神经解剖显示，受神经支配的鸡趾大小与在合适的位置支配鸡趾的神经元数目相吻合，即移植的鸡趾越大，支配它的神经元数目就越多。进一步研究还表明，移植多余的鸡趾不仅能够增加支配它的神经元数目，还能增加位于其他位置并非支配它的神经元的数目。因此，她认为可能有一种扩散性的分泌因子，能够在一定的距离范围内增加神经元数目。很快她设计了一个巧妙的实验证明了这个设想：她将某些组织添加到鸡胚尿囊膜（chick embryo chorioallantoic membrane）的上方，结果能增加正在发育中的胚胎内的神经元数目。鸡胚尿囊膜是将胚胎与鸡蛋气袋分开的一层膜，它能让气体和小分子物质通透，但细胞不行。这说明在组织液中一定存在某种可以通过鸡胚尿囊膜的因子，从添加的组织扩散到鸡胚内，控制神经元的命运，于是 Levi-Montalcini 决定开始纯化这种神秘的因子。

然而在当时，对于这种因子的化学本质还一无所知，而且也不知道它真正的生理功能是否就是提高神经元的数目。只是到了后来，才证实它就是一种 NGF，其功能是在胚胎发育的晚期控制神经细胞死亡的数目，从而导致神经元数目增加和神经节（ganglion）增大。

就像纯化其他任何一种蛋白质一样，在纯化 NGF 之前，必须首先建立一种有效的测活方法。但就当时的研究水平，研究者只能在对鸡胚进行神经解剖以后，通过测量神经节的大小和存活在神经节的神经元数目来测活。而在正在发育的鸡胚中，进行这样的解剖显然既耗时又烦琐。如果以此作为测定 NGF 活性的基础，那么整个纯化工作的进展将会极其缓慢。

Levi-Montalcini 在意识到使用上述鸡胚测活系统的困难以后，决定去巴西大学的 Carlos Chagus 实验室学习神经组织培养技术，以改进她的测活方法。当时的 Carlos Chagus 实验室已成功地从幼小动物分离出背根神经节（dorsal root ganglion，DRG），并能够让它们在培养液中存活。而对已分离的 DRG 的解剖要简单得多。作为一个访问学者，Levi-Montalcini 有一个惊人的发现：将带有一种肿瘤组织（代号为 S-180）的 DRG 在体外进行培养，DRG 的形态发生了很显著的变化。经过分析发现，原来是从构成 DRG 的神经元上伸出了很多突触。这个发现是一个重要的突破，它不仅显示了某种扩散性分子的存在，还提供了一种很方便的用于纯化该分子的测活手段。

使用 DRG 测活非常重要。首先，这种方法很方便，易于快速测定，但测定的生物活性跟最初的生物活性已有微妙的变化（测定突触的增加，而不是原来的神经元数目的增加）。这种测活手段后来成为寻找新的神经生长因子的一种标准。也许这样的因子可能（也可能不）会提高神经元的数目。事实最终证明，NGF 具有双重活性。因而活性测定方法的改变没有影响到对 NGF 本身的纯化。但如果参与神经伸展和促进神经元存活的不是同一种物质，那么 Levi-Montalcini 纯化到的将是另外一种因子。Levi-Montalcini 是幸运的，因为 NGF 同时具有两种活性。

有了这种测活方法，Levi-Montalcini 开始尝试从 S-180 组织中纯化 NGF。但不幸的是，S-180 中的 NGF 浓度很低，利用当时的纯化技术，活性很容易损失，因此并没有取得成功。然而，就在同时期，人们发现，在蛋白质纯化中使用蛋白酶部分水解组织，有助于蛋白质的溶解。在当时，蛇毒经常被作为蛋白酶很好的来源，于是她用蛇毒对 S-180 的抽取物进行了处理。结果她意外地发现，蛇毒处理能够导致抽取物中

NGF 生物活性的显著提高。对照实验表明，升高的生物活性是从蛇毒中得到的，而不是来自 S-180。这个发现促使 Levi-Montalcini 考虑使用蛇毒来代替 S-180 组织作为分离 NGF 的材料。

蛇毒的确是 NGF 很好的来源，但获取大量的蛇毒并非易事。于是，Levi-Montalcini 想到从其他与分泌蛇毒的腺体有关的组织中寻找 NGF。她很快发现，小鼠唾液腺是 NGF 的丰富来源，并且它比蛇毒更容易获取。于是，纯化 NGF 的原材料又转到了唾液腺。

一个有趣的问题是，存在于小鼠唾液腺中的 NGF 就是原来一直想从 S-180 组织中得到的 NGF 吗？在 20 世纪 80 年代，John Wagner 和 Pat D'Amore 已经发现，一种称为酸性成纤维细胞生长因子（fibroblast growth factor，FGF）的蛋白质能够导致 PC12 细胞（一种神经元分化细胞株）的突触伸展。那么，是不是 S-180 产生的不是 NGF 而是酸性 FGF 呢？于是有人使用肝素亲和层析柱对 S-180 抽取物进行分离，结果得到大小相同的两个峰，第一个峰洗脱的位置与酸性 FGF 重叠，另一个刚好在 NGF 的位置！面对这样的结果，我们真应该为 Levi-Montalcini 感到庆幸：如果她一直使用 S-180 作为分离 NGF 的来源，而且只收集第一个峰的话，那么最终被她纯化到的蛋白质一定是酸性的 FGF。

如果阅读几篇关于 NGF 纯化的早期论文，你会发现那时的纯化需要很多步骤，似乎有点劳民伤财，但正是这些早期工作使得 NGF 的结构最终得到确定。而随着人们对 NGF 结构理解的深入，研究者逐渐设计出更快速、更有效的纯化方案。Mobley 建立的两步快速纯化程序就是基于唾液腺最初分泌出来的 NGF 和其他几种蛋白质分子形成复合物的信息。该复合物被称为 7S NGF。如果将从唾液腺制备的抽取物的杂质除去，直接通过羧甲基纤维素柱，几乎没有任何杂蛋白能被除去，于是就没有 NGF 的富集。但如果调低这种柱的流出液的 pH，则发现某些物质开始沉淀，7S NGF 发生解离，形成 β-NGF。如果将溶液调回到原来的第一次柱层析中使用的 pH 和盐浓度以后，再上第二根柱，那么 β-NGF 就被吸附在柱子上，而大多数杂蛋白直接流出。再经过高 pH 和高盐溶液洗脱后，NGF 被洗脱下来，其纯度可达 99%。

Levi-Montalcini 是幸运的，她与发现表皮生长因子（EGF）的 Stanley Cohen 同时获得 1986 年的诺贝尔生理学或医学奖。

e4-5 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Protein_purification（维基百科有关蛋白质纯化的内容）
2. http://www.basic.northwestern.edu/biotools/proteincalc.html（美国西北大学提供的一个在线计算肽段性质的计算器）
3. http://www.protocol-online.org/prot/Molecular_Biology/Protein/index.html（在线提供各种生物化学、细胞生物学和分子生物学等学科的研究方法）
4. http://expasy.hcuge.ch（一个生物信息学门户网站，内有双向电泳数据库）

参考文献：

1. Mondal K, et al. Affinity based strategies for protein purification. *Anal Chem*, 78: 3 499-3 504, 2006.
2. Labrou N E. Design and selection of ligands for affinity chromatography. *J Chromatogr B*, 790: 67-78, 2003.
3. Bocchiniv A. The nerve growth factor: purification as a 30 000-molecular-weight protein. *Proc Natl Acad Sci USA*, 64: 787, 1969.
4. Varon S, et al. The isolation of the mouse nerve growth factor protein in a high molecular weight form. *Biochemistry*, 6: 2 202, 1967.

第五章　核苷酸

核苷酸(nucleotide)是一类极为重要的生物小分子,在细胞内参与多项重要的生物功能,其最重要的功能无疑是作为核酸的基本组成单位,为生物编码遗传信息。本章将重点介绍核苷酸的化学结构、理化性质及其生物学功能。

第一节　核苷酸的结构与组成

核苷酸由核苷(nucleoside)和无机磷酸基团组成。核苷实为一种戊糖苷(pentoside),由碱基与 D-核糖或 D- 脱氧核糖通过 β-N 糖苷键连接而成。

一、碱基

碱基是核苷酸中最重要的部分,因为在核酸分子中,编码遗传信息的是特定的碱基序列。

(一) 碱基的结构

碱基也被称为含氮碱基(nitrogenous base),它们是含有 N 原子的嘌呤(purine)或嘧啶(pyrimidine)的衍生物。衍生于嘧啶的碱基称为嘧啶碱基,衍生于嘌呤的碱基称为嘌呤碱基。

嘧啶是一种六元的芳香杂环,含有 2 个 N 原子,为一个平面结构;嘌呤由六元的嘧啶环与五元的咪唑环(imidazole ring)融合而成,环上共有 9 个原子,其中有 4 个 N 原子。在嘧啶环和咪唑环之间有小的弯曲,故嘌呤环不完全在一个平面上。嘧啶和嘌呤按 IUPAC(国际纯粹与应用化学联合会)命名法的原子编号见图 5-1。

e5-1 嘧啶环和嘌呤环编号记忆之秘诀

(1)

嘧啶环　胞嘧啶 (2-氧-4-氨基嘧啶)　尿嘧啶 (2-氧-4-氧嘧啶)　胸腺嘧啶 (2-氧-4-氧-5-甲基嘧啶)

(2)

嘌呤环　腺嘌呤 (6-氨基嘌呤)　鸟嘌呤 (2-氨基-6-氧嘌呤)　次黄嘌呤 (6-氧嘌呤)　黄嘌呤 (2-氧-6-氧嘌呤)

图 5-1　嘧啶环和嘌呤环的编号以及各种碱基的化学结构

(二) 常见的碱基

生物体内最常见的嘧啶碱基是胞嘧啶(cytosine,C)、尿嘧啶(uracil,U)和胸腺嘧啶(thymine,T),胸腺嘧啶即是 5- 甲基尿嘧啶(5-methyluracil);嘌呤碱基是腺嘌呤(adenine,A)和鸟嘌呤(guanine,G)(图 5-1)。RNA 和 DNA 共有的碱基是胞嘧啶、腺嘌呤和鸟嘌呤,而尿嘧啶通常只存在于 RNA,胸腺嘧啶通

常只存在于 DNA。但有时 DNA 分子中也会有少量的 U,某些 RNA 分子还会有少量的 T(如 tRNA)。

(三) 修饰碱基

除了以上 5 种常见的碱基以外,机体内还存在着上百种以上的修饰碱基,这些修饰碱基有的很少见,但多为 5 种常见碱基的修饰产物或代谢产物,例如 5- 甲基胞嘧啶、5- 羟甲基胞嘧啶、次黄嘌呤(hypoxanthine)、黄嘌呤(xanthine)、尿酸(uric acid)和二氢尿嘧啶(dihydrouracil)。这些修饰碱基中,5-甲基胞嘧啶特别重要,如果它出现在真核生物的基因组 DNA 上,可抑制特定基因的表达,从而在一个基因序列不变的情况下改变生物的表型。因此,5- 甲基胞嘧啶有时被视为 DNA 分子上的第五个碱基。茶叶中的有效成分——茶碱(theobromine)和咖啡中的有效成分——咖啡因(caffeine)都属于嘌呤碱基(图 5-2)。这两种天然的嘌呤碱基在体内能够加强肾上腺素的作用,因而具有兴奋之效(参看第十七章“激素及其受体介导的信号转导”)。

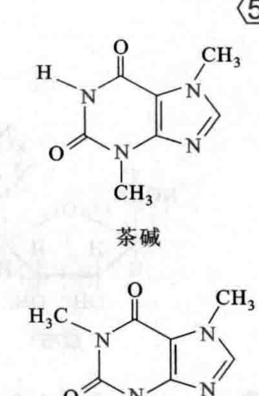


图 5-2 茶碱和咖啡因的化学结构

(四) 碱基的性质

1. 水溶性差

这与其芳香族的疏水性杂环结构有关。碱基的疏水性质对于 DNA 形成稳定的双螺旋结构非常重要。

(1)

酮式(99.99%) 烯醇式(0.01%)

(2)

氨基式(99.99%) 亚氨基式(0.01%)

图 5-3 碱基的互变异构

2. 互变异构

嘧啶环和嘌呤环的芳香族性质以及环上取代基团(羟基和氨基)的富电子性质致使它们在溶液中,能够发生酮式(keto)- 烯醇式(enol)或氨基式 - 亚氨基式的互变异构(tautomeric shift)(图 5-3)。虽然碱基的这两种异构体是可以相互转变的,但在体内的主要形式为更稳定的酮式。这一点对于在 DNA 双螺旋结构中形成正确的碱基配对十分关键。

3. 酸碱解离

碱基上含有可解离的酸碱基团,这些基团的 p*K*a 不一样。在中性 pH 下,碱基主要以内酰胺形式存在。尿嘧啶的 N1 和 N3 的 pK_a 均超过 8;与此相比,胞嘧啶的 N3 的 pK_a 为 4.5。各个基团的 pK_a 决定了在生理 pH 下质子是否与环上的各个 N 原子结合,而结合与否又决定了这些 N 原子能作为氢键的供体还是受体。究竟是作为氢键受体还是供体将直接决定核酸分子中碱基配对的性质,这是双螺旋结构形成的基础,对于核酸的生物学功能至关重要(参看下一章“核酸的结构与功能”)。

Quiz1 各个碱基的 pK_a 与各个氨基酸残基的 pK_a 有何差别?这种差别对核酸和蛋白质的功能带来什么影响?

4. 强烈的紫外吸收

碱基杂环上的共轭双键造成碱基对于紫外线具有强烈的吸收能力,其最大吸收值在 260 nm。此性质可用来定性或定量测定碱基和含有碱基的化合物(核苷酸和核酸)。

Quiz2 嘧啶碱基和嘌呤碱基相比,哪一种的紫外吸收能力更强?

二、核苷

(一) 核苷的结构

核苷(nucleoside)是由戊糖和碱基通过 *β-N-* 糖苷键形成的糖苷,糖苷键由戊糖的异头体 C 原子与嘧啶碱基的 N1 或嘌呤碱基的 N9 形成(图 5-4)。核苷中的戊糖有 D- 核糖和 2- 脱氧 -D- 核糖两种,它们都以呋喃型环状结构存在(参看第十五章“糖类”)。其中由核糖形成的核苷叫做核糖核苷

腺苷　鸟苷　胞苷　尿苷

图 5-4　四种嘌呤核苷和嘧啶核苷的化学结构

(ribonucleoside),由脱氧核糖形成的核苷叫做脱氧核苷(deoxyribonucleoside)。为了避免呋喃糖环与碱基环在原子的编号上出现混淆,需要在呋喃环上各原子编号的阿拉伯数字后加“′”。

Quiz3 “′”在英文中怎么读?

核糖上的 2′-羟基是否存在似乎是很小的差别,但它对于核酸的二级结构以及核酸对碱水解和酶促水解的敏感性则会产生巨大的影响(参看第七章“核酸的性质及研究方法”)。

(二) 核苷的构象

在核苷中,碱基在糖苷键上的旋转受到空间位阻(特别是 C2′上的 H)的限制。结果核苷和核苷酸能以两种构象存在,即顺式(syn)和反式(anti)(图 5-5)。顺式核苷的碱基与戊糖环在同一个方向,反式核苷的碱基与戊糖环在相反的方向。

由于嘧啶环 O2 和戊糖环 C5′之间的空间位阻,嘧啶核苷的构象通常为反式。嘌呤核苷可采取两种构象,但无论是哪一种,大致呈平面的呋喃糖环和碱基环不是共平面的,而是相互间近似垂直。自由的嘌呤核苷(特别是鸟苷)更容易形成顺式构象,但在 DNA 或 RNA 双螺旋中,嘌呤核苷主要以反式构象存在。

顺式腺苷　反式腺苷　反式尿苷

图 5-5　顺式核苷和反式核苷

(三) 常见的核苷

常见的核苷由常见的碱基与核糖或脱氧核糖组成。核糖核苷有腺苷(adenosine)、鸟苷(guanosine)、胞苷(cytodine)和尿苷(uridine);脱氧核糖核苷有脱氧腺苷(deoxyadenosine)、脱氧鸟苷(deoxyguanosine)、脱氧胞苷(deoxycytodine)和脱氧胸苷(thymidine)。

(四) 修饰核苷

修饰核苷主要是指由修饰碱基与戊糖组成的核苷,也包括核糖环被修饰的核苷以及少数不是以 N-糖苷键相连的核苷,例如在假尿苷(pseudo-uridine, ψ)中,尿嘧啶碱基和核糖以 C—C 键相连(图 5-6)。修饰核苷通常存在于 RNA 分子上,尤其是 tRNA 和 rRNA。

(五) 核苷的性质

核苷的水溶性要比自由的碱基高得多,这与核糖基的高度亲水性有关。与糖苷一样,核苷在碱性条件下较稳定。嘧啶核苷还能抵抗酸水解,但嘌呤核苷很容易发生酸水解。

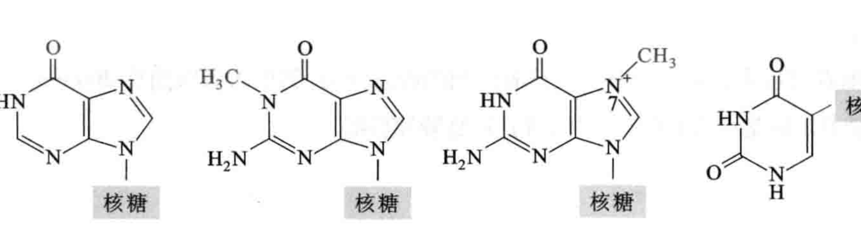


图 5-6　5 种修饰的核苷

三、核苷酸

（一）核苷酸的结构

核苷酸是核苷的戊糖羟基发生磷酸化反应而形成的磷酸酯。其中核糖核苷的磷酸酯为核糖核苷酸(ribonucleotide),脱氧核苷的磷酸酯为脱氧核苷酸(deoxyribonucleotide)。理论上,核苷的5′-OH、3′-OH和2′-OH均可以发生磷酸化,而分别形成核苷-5′-磷酸、核苷-3′-磷酸和核苷-2′-磷酸。但是,自然界的核苷酸多为核苷-5′-磷酸。常见的核苷酸与核苷的名称和结构见表5-1和图5-7。

表5-1 常见的核苷和核苷酸

核糖核苷	腺苷	鸟苷	胞苷	尿苷
脱氧核苷	脱氧腺苷	脱氧鸟苷	脱氧胞苷	脱氧胸苷
核糖核苷酸	腺苷酸(AMP)	鸟苷酸(GMP)	胞苷酸(CMP)	尿苷酸(UMP)
脱氧核苷酸	脱氧腺苷酸(dAMP)	脱氧鸟苷酸(dGMP)	脱氧胞苷酸(dCMP)	脱氧胸苷酸(dTMP)

图5-7 常见的5种核糖核苷酸的化学结构

由于自然界游离的核苷酸多为核苷-5′-磷酸,因此如果没有特别说明,某某核苷酸即指核苷-5′-磷酸。

（二）核苷单磷酸、核苷二磷酸和核苷三磷酸

核苷单磷酸(nucleoside monophosphate,NMP)是指核苷的单磷酸酯。核苷单磷酸还可以通过两次成酐反应,分别形成核苷二磷酸(nucleoside diphosphate,NDP)和核苷三磷酸(nucleoside triphosphate,NTP)(图5-8)。为了将核苷二磷酸和核苷三磷酸上不同的磷酸根区分开来,将直接与戊糖5′-羟基相连的磷酸基团定为α磷酸,其余两个磷酸基团从里到外依次称为β磷酸和γ磷酸。

（三）核苷酸的英文缩写

通常用英文缩写表示各种形式的核苷酸,以NMP(rNMP)、NDP(rNDP)和NTP(rNTP)分别表示核糖核苷单磷酸、核糖核苷二磷酸和核糖核苷三磷酸,r可以被省掉;dNMP、dNDP和dNTP分别表示脱氧核苷单磷酸、脱氧核苷二磷酸和脱氧核苷三磷酸。遇到具体的核苷酸,使用碱基首字母代替N。例如ATP和dATP分别表示腺苷三磷酸和脱氧腺苷三磷酸,CDP和dCMP分别表示胞苷二磷酸和脱氧

图5-8　腺苷单磷酸→腺苷二磷酸→腺苷三磷酸的转变

胞苷单磷酸,IMP 表示次黄苷单磷酸。

（四）环核苷酸

某些核苷三磷酸在特定的条件下,受环化酶的催化,形成环核苷酸,在细胞中作为第二信使,例如3′,5′-环腺苷酸（cyclic AMP,cAMP）、3′,5′-环鸟苷酸（cyclic GMP,cGMP）和环二鸟苷酸（cyclic diguanylate,c-di-GMP）（图5-9）。其中cAMP既可以在细菌体内作为一种第二信使,激活多种与糖类分解代谢有关的基因表达（参看第四十二章"原核生物的基因表达调控"）;也可以在高等生物体内与cGMP一样,作为多种激素的第二信使（参看第十七章"激素及其受体介导的信号转导"）。c-di-GMP则在多种细菌中作为第二信使,参与调节生物薄膜（biofilm）的形成和致病因子的产生等。

图5-9　环腺苷酸、环鸟苷酸和环二鸟苷酸的化学结构

（五）核苷酸的构象

核苷酸的构象不是固定不变的，它也具有一定的柔性。具体表现如下：

（1）核苷酸分子上 C–O–P 内的单键可以旋转。

（2）一个核苷酸具有 7 个扭角，即 α、β、γ、δ、ε、ζ 和 χ。其中 χ 在碱基和核糖之间，其值受到很大的限制（图 5–10）。

（3）脱氧核糖环不是平坦的，而是形成一定的折皱，共有 3′– 内（3′–endo）、2′– 内（2′–endo）、3′– 内 /2′– 外（3′–endo/2′–exo）和 2′– 内 /3′– 外（2′–endo/3′–exo）等 4 种形式（图 5–11）。不同的折皱影响到 5′–OH、3′–OH 上磷酸根的位置。

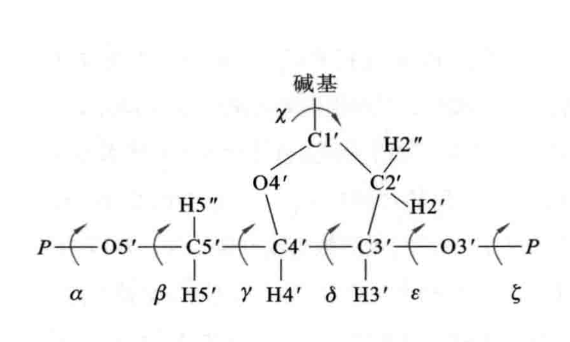


图 5–10　核苷酸的 7 个扭角

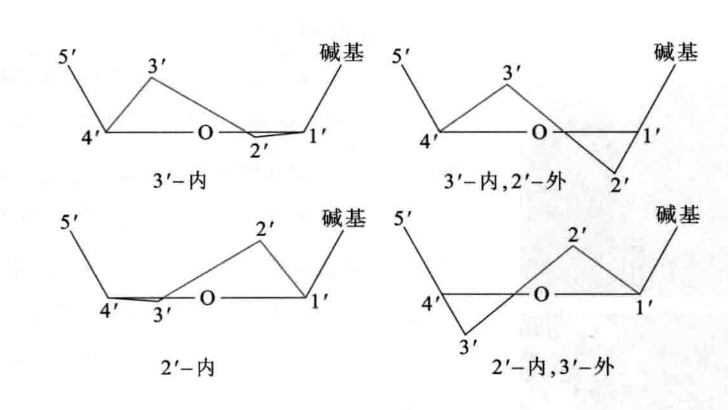


图 5–11　脱氧核糖环的折皱

（4）N 糖苷键有顺式和反式（详见核苷的结构）。

（六）核苷酸的理化性质

核苷酸的某些性质由碱基决定，例如紫外吸收、互变异构；某些性质由核糖或脱氧核糖决定，例如易溶于水和具有旋光性；某些性质由磷酸基团和碱基共同决定的，例如核苷酸的两性解离和具有等电点；某些性质与 N 糖苷键有关，如核苷酸特别是嘌呤核苷酸在酸性溶液中不稳定，易发生脱碱基反应。

第二节　核苷酸的功能

核苷酸的生物功能主要包括以下几个方面：

（1）作为核酸合成的前体，其中 NTP 和 dNTP 分别是 RNA 和 DNA 合成的前体。

（2）充当能量货币，驱动细胞内各种需能反应。其中，ATP 是细胞最通用的能量货币，有时细胞也使用 UTP（糖原合成时所需要）、CTP（磷脂合成时所需要）和 GTP（蛋白质合成时所需要）。

（3）参与细胞的信号转导。例如，cAMP 和 cGMP 作为某些激素的第二信使，鸟苷酸能够调节 G 蛋白的活性（参看第十七章“激素及其受体介导的信号转导”）。

（4）作为其他物质的前体或辅酶 / 辅基的成分，如 ADP 为辅酶 Ⅰ 和 Ⅱ 的组分，鸟苷酸作为 Ⅰ 组内含子（group Ⅰ intron）核酶的辅酶。

（5）转变成一些特殊的活化中间物，参与体内某些物质的合成。例如，UDP – 葡萄糖和 CDP – 乙醇胺分别参与糖原和磷脂酰乙醇胺的合成；S – 腺苷甲硫氨酸（S–adenosylmethionine，SAM）作为甲基供体，参与 RNA 和 DNA 的甲基化修饰反应。

（6）作为某些酶的别构效应物参与代谢的调节。例如，ATP 为磷酸果糖激酶 –1 的负别构效应物抑制糖酵解，AMP 作为糖原磷酸化酶的正别构效应物激活糖原的分解（参见第二十二章“糖酵解”和第二十七章“糖原代谢”）。

框 5-1　生化热点——生命系统中的砷可以完全取代磷吗?

一直以来,我们对组成生命物质基本元素的认识都是六大元素——C、H、O、N、S、P。我们也一直认为,磷在生命活动中的重要性几乎是不可替代的。生命系统普遍使用的通用“能量货币”——ATP绝对需要磷元素;生命的遗传物质DNA就是靠磷酸二酯键把无数个核苷酸串联起来的;细胞内的各种信号转导通路最后基本上都是靠蛋白质的磷酸化修饰来起作用的;还有,组成生物膜的重要成分磷脂也需要磷。

然而,2010年12月2日,Felisa Wolfe-Simon在*Science*上发表的一篇引起广泛争议的题为“一种细菌能在用砷取代了磷的环境下生长”(A Bacterium That Can Grow by Using Arsenic Instead of Phosphorus)的论文可以说颠覆了这一传统认识。



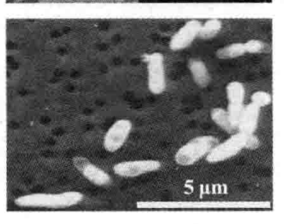
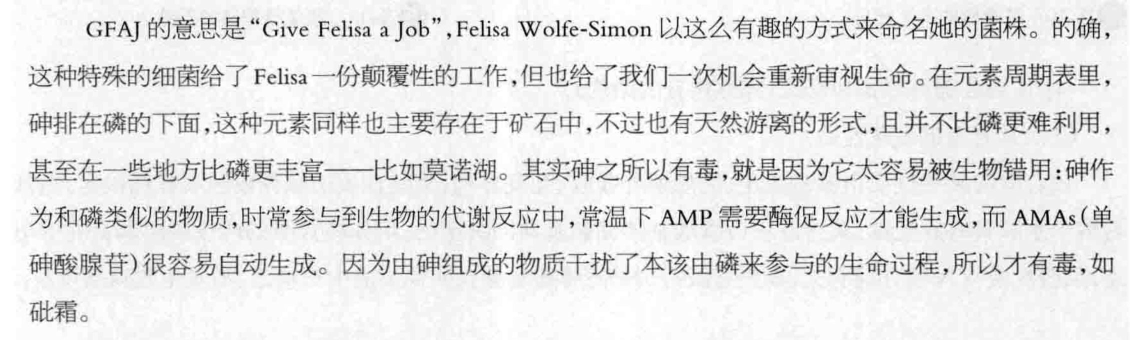


图 5-12　Mono湖边的Felisa Wolfe-Simon和显微镜下的GFAJ-1

*Science*上报道的这项由美国NASA资助、由航天生物研究所Felisa主持的研究发现,在一些极端环境下,个别微生物能够利用砷取代磷的所有功能。他们从美国加利福尼亚州的莫诺湖(Mono lake)内,成功分离出一种单胞体细菌(***Halomonadaceae***),该湖内的湖水盐度和碱度极高,也是世界上天然水体里砷含量最高的地方之一,这种极端的生境造就了不少耐受极端环境的生物。而Felisa在湖底发现并在实验室成功培育的GFAJ-1菌株,比单纯耐受环境更向前进了一步,对大多数生物有毒的砷,能被它用来作为磷的替代物成为细胞的基本元素,即完全可以利用砷取代磷在生命活动中扮演很多重要功能。放射性同位素标记证实了GFAJ-1用砷组成自身,因此这株细菌在某种程度上说是一种“砷基生命”(图5-12)。



GFAJ的意思是“Give Felisa a Job”,Felisa Wolfe-Simon以这么有趣的方式来命名她的菌株。的确,这种特殊的细菌给了Felisa一份颠覆性的工作,但也给了我们一次机会重新审视生命。在元素周期表里,砷排在磷的下面,这种元素同样也主要存在于矿石中,不过也有天然游离的形式,且并不比磷更难利用,甚至在一些地方比磷更丰富——比如莫诺湖。其实砷之所以有毒,就是因为它太容易被生物错用:砷作为和磷类似的物质,时常参与到生物的代谢反应中,常温下AMP需要酶促反应才能生成,而AMAs(单砷酸腺苷)很容易自动生成。因为由砷组成的物质干扰了本该由磷来参与的生命过程,所以才有毒,如砒霜。

通常条件下,生物体并不能接受砷作为生命的一大元素。原因是常温下,砷的化合物过于活泼,不如磷化合物稳定,如果DNA分子里面有砷酸用作连接物,DNA链就很容易从砷酸那里断裂。可是,GFAJ-1菌株,作为一种罕见的嗜极细菌,做到了把砷安全地利用起来:它也许并没有把细胞中的磷全部替换为砷,但它们含有砷酸的DNA甚至经受了实验室提取的一系列过程的考验,在对基因组进行详细分析之前,暂时还不知道它们是怎么做到的。

其实,在莫诺湖发现利用砷的生物已经不是新闻。早在2008年,就有人发现了用光能氧化亚砷酸,而不是光解水来进行光合作用的细菌。不过这种细菌只是将砷的化合物作为光反应中的电子供体来固定二氧化碳,但是这次发现的GFAJ-1菌株却在用砷组成自身的各种关键分子,因此被称作“砷基生命”而引起了学术界的轰动。GFAJ-1的不同之处在于,它第一个证实了生命必需的元素可以被替代的想法,因此,那些过去被描述为“不适合生命”的外星环境,也许需要重新审视,它们也许对于那个地方的“土著生命”来说是一片沃土。

科学家们推测这些以砷为生的生命说不定就是我们这颗星球上很古老的居民。在地球上的磷还没有被生物圈搬运到全世界的时候,也许很多地方的生物只能靠砷吃饭,用亚砷酸来进行光合作用的生物也许也比现代我们熟悉的光合有机体出现得早。如果原始地球是这样一种景象,那么外星球呢?人们认为砷基生命的结构不稳定,只是在地球上的温度条件下,而在-180℃的土卫六的情况又如何呢?

那些“不适合生物生存”的星球也许只是由于我们认识局限的一厢情愿，至少在今天，我们知道生物学家们长期以来熟悉的地球生命六大元素就有可能被取代。“科学始于当你不相信专家之时。”NASA 发布会上的一位嘉宾这样说。

然而，在 Felisa 的论文发表一年以后，加拿大微生物学家 Rosie Redfield 在 2012 年 1 月 20 日声称，以她为首的研究小组没有成功地重复出 Felisa 的实验结果。本人也曾经冒昧地给 Felisa 发过 email，询问 GFAJ-1 的进展，但没有得到她的回复。目前有关 GFAJ-1 的争议还在继续进行之中，让我们耐心地等待着最终的定论！

（7）调节基因表达。例如，细菌在氨基酸饥饿的时候，产生鸟苷 -5′- 二磷酸 -3′- 二磷酸（ppGpp）和鸟苷 -5′- 三磷酸 -3′- 二磷酸（pppGpp），以此来调节基因的表达，诱发大肠杆菌的严紧反应（stringent response）（参看第四十二章“原核生物的基因表达调控”）。

e5-2 本章小结填填看

推荐网址：

1. http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/N/Nucleotides.html（内有核苷酸相关内容，简单实用）
2. http://www.basic.northwestern.edu/biotools/OligoCalc.htm（美国西北大学提供的一个在线计算寡核苷酸性质的计算器）

参考文献：

1. Wolfe-Simon F, et al. A bacterium that can grow by using arsenic instead of phosphorus. *Science*, 332: 1 163–1 166, 2011.
2. Miller S I, et al. Cyclic di–GMP as a bacterial second messenger. *Microbiology*, 150: 2 497–2 502, 2004.

第六章　核酸的结构与功能

核酸即多聚核苷酸，是由多个核苷酸通过 3′，5′－磷酸二酯键相连的多聚物，可分为核糖核酸（RNA）和脱氧核糖核酸（DNA）两类，其中 DNA 是生物体主要的遗传物质，只有少数病毒以 RNA 为遗传物质。DNA 能作为遗传物质与其结构是分不开的，其一级结构贮存遗传信息，二级结构有助于遗传物质的复制、转录、重组和修复。与 DNA 单一的功能相比，RNA 则是一个多面手，其功能是多方面的，除了作为某些病毒的遗传物质，它还可以在基因表达中起主导作用，甚至还可以作为酶起催化作用。RNA 功能的多样性与其复杂多变的结构有关。

本章将重点介绍核酸的一级结构、二级结构和三级结构，特别是 DNA 的二级结构，此外，还会对核酸和蛋白质形成的复合物做一般介绍。

第一节　核酸的分类

核酸的基本组成单位是单核苷酸。构成 DNA 和 RNA 的单核苷酸分别是脱氧核苷酸和核糖核苷酸。

核糖核苷酸含有 2′-OH，因此在理论上，多聚核糖核苷酸还有另外一种连接方式，即 2′，5′－磷酸二酯键连接。然而，在自然界很难找到以这种方式连接的多聚核苷酸。不过，高等动物体内的干扰素（interferon）在作用的时候，可诱导靶细胞合成一种以 2′，5′－磷酸二酯键连接的多聚腺苷酸。

RNA 与 DNA 主要有三大差别（表 6-1）：

► 表 6-1　DNA 和 RNA 的结构异同

性质	RNA	DNA
戊糖	D- 核糖	2′-D- 脱氧核糖
碱基	第四个碱基通常是 U	第四个碱基通常是 T
多聚核苷酸链的数目	多为单链	多为双链
在细胞内的双螺旋	A 型	B 型和 Z 型
种类	多种	只有一种
功能	功能多样	只有一种功能，即充当遗传物质
碱溶液下的稳定性	不稳定，很容易水解	稳定

(1) RNA 分子中的戊糖是核糖，而 DNA 分子中的戊糖是脱氧核糖　这项差别对于两类核酸各自的结构与功能具有重大的影响，也是我们判断一种核酸是 RNA 还是 DNA 的唯一标准。DNA 因为是脱氧核糖，缺乏反应性的亲核基团——2′-OH，降低了磷酸二酯键自发水解的速率，提高了 DNA 的稳定性，这对于遗传物质至关重要。RNA 中的核糖带有 2′-OH，2′-OH 的亲核性使其很不稳定，容易发生水解，特别是在碱性溶液里。这使得 RNA 并不适合充当遗传物质，但却适合在细胞里充当蛋白质合成的模板（mRNA）——在需要的时候被转录出来，在不需要的时候可迅速降解。当然，在细胞内有多种不同类型的 RNA，它们的稳定性是不一样的，像 tRNA 和 rRNA 就特别稳定。tRNA 和 rRNA 之所以特别稳定，原因一方面是它们在细胞里不断地被循环使用，细胞需要它们稳定；另一方面是它们的很多 2′-OH 发生了甲基化修饰，而大大降低了 2′-O 的反应性。RNA 带有 2′-OH 还有一个用处，就是使其能够利用羟基的亲核进攻性，作为酶催化一些重要的生化反应（参看第十一章“核酶的结构与

功能”)。

(2) RNA 的第四个碱基通常是 U,而 DNA 的第四个碱基通常是 T　这项差别不是绝对的,因为有的 RNA 分子中也含有少量的 T,而 DNA 分子中经常含有少量的 U。RNA 分子中的 T 是由转录出来的 U 发生甲基化修饰产生的。DNA 分子上的 U 则有两种机制可以产生:一是因为细胞中有少量的 dUTP,它在 DNA 复制的时候,有可能代替 T 而直接参入到新合成的 DNA 链上,二是 DNA 分子上的 C 自发地脱氨基而变成 U。无论是哪一种途径产生的 U,只要出现在 DNA 分子上,就会被细胞内的 DNA 修复系统视为损伤而很快将其修复——在复制中参入的 U 会被修复成 T,由 C 变来的 U 则被修复成 C(参看第三十四章“DNA 损伤、修复和突变”)。显然要让修复系统能正确地行使功能,这些“坏 U”必须能够被及时发现。如果 DNA 分子中的第四个碱基也是 U,那这些 U 就属于“好 U”。在一个 DNA 分子上同时出现“坏 U”和“好 U”的情况下,修复系统是区分不了的,更谈不上修复了。因此,DNA 分子中的第四个碱基用 T 代替 U,等于是打上一种标记,可让修复系统能够及时发现那些错误的 U 并将它们修复,从而有效防止了 C → U 的突变。

框 6-1　生化聚焦——DNA 分子一定没有“你”(U)吗?

DNA 分子上一般是不允许有 U 的,若是出现 U,细胞内的修复系统会将其视为损伤,迟早将其切除,然后换成与互补链碱基相配对的碱基。U 如果是由 C 脱氨基变来的,就换成 C;U 如果是在 DNA 复制中直接参入的,就换成 T。但事实并非总是如此,迄今为止,已发现某些噬菌体的 DNA 完全用 U 代替 T,此外,某些生物在发育的某个阶段,会故意让细胞内 DNA 在复制的时候,参入更多的 U,这两种情况的出现对于相关生物肯定具有某种进化上的优势。

在 DNA 基因组上完全用 U 取代 T 的一个典型例子就是枯草杆菌的 PBS2 噬菌体,这种噬菌体为了让它的基因组 DNA 能稳定地参入 U,一方面通过自身编码的一些酶,改造宿主细胞内的核苷酸合成代谢途径,以提高宿主细胞内 dUTP 的量,另一方面是编码一种蛋白质抑制剂,专门抑制宿主细胞内用来切除 DNA 分子上 U 的尿嘧啶 -DNA 糖苷酶。那么,PBS2 噬菌体用 U 代替 T 有什么好处呢?这应该是有利于保护它的基因组 DNA,使其免受宿主细胞编码的限制性内切酶的水解。

含有 U 的 DNA 还可能在完全变态昆虫的发育过程中起重要作用。这些昆虫(如蝴蝶)在生活周期经历化蛹,它们缺乏编码尿嘧啶 -DNA 糖苷酶的主要基因,因此无法去除 DNA 分子上可能出现的 U。在果蝇的幼虫阶段,那些成体不需要的组织细胞降解 dUTP 和合成 dTTP 前体的酶水平极低,导致 [dUTP]/[dTTP] 的比值异常高,结果这些组织中的 DNA 在复制的时候,会参入大量的 U。由于细胞内本来就缺乏尿嘧啶 -DNA 糖苷酶,因此无法修复 DNA 中的 U,含有大量 U 的 DNA 反而作为一种细胞死亡的信号,使得有大量尿嘧啶 -DNA 的组织在化蛹阶段被选择性降解。

尿嘧啶在高等动物抗体多样性的产生方面也有一定贡献,抗体多样性的产生对于动物对外来各种抗原的防御十分重要,有多种机制可产生抗体的多样性,其中一种就是编码抗体的基因在 B 淋巴细胞前体内进行超突变(hypermutation)。超突变开始于一种特定的胞嘧啶脱氨酶,它催化 DNA 分子上特定区域发生 C 脱氨基变成 U 的反应。通过这种方式,激发 B 淋巴细胞前体内的易错修复系统,从而将新的序列引入到编码抗体的基因片段上。当然,这种系统是受到严格调控的,以防止失控而导致细胞癌变。

(3) RNA 通常是单链的,DNA 通常是双链的　这项差别也不是绝对的,原因是有的 RNA 是双链的,比如某些 RNA 病毒和真核细胞内的 miRNA,也有的 DNA 是单链的,例如某些 DNA 病毒。RNA 处于单链状态,使其能够自我折叠成可以和蛋白质相媲美的各种类型的二级结构和三级结构,这是形成 RNA 结构多样性的基础,否则所有的 RNA 与 DNA 一样,只能形成千篇一律的双螺旋。RNA 三维结构的多样性使其在细胞内能行使多项生物学功能。DNA 通常是双链的,使其能

够充分地行使作为遗传物质这项唯一的功能(参看本章第三节有关 DNA 双螺旋结构发现的意义的描述)。

已在生物体内发现多种天然的 RNA,例如转移 RNA(tRNA)、信使 RNA(mRNA)、核糖体 RNA(rRNA)、核小 RNA(small nuclear RNA,snRNA)、核仁小 RNA(small nucleolar RNA,snoRNA)、微 RNA(microRNA,miRNA)、小干扰 RNA(small interfering RNA,siRNA)、小激活 RNA(small activating RNA,saRNA)、长间插非编码 RNA(long intervening non-coding RNA,lincRNA)、7SL RNA、向导 RNA(guide RNA,gRNA)和 Xist RNA 等。这些 RNA 具有特殊的结构和功能,其中某些 RNA 存在于所有的生物中,某些 RNA 是真核生物或原核生物特有的。根据 RNA 是否具有编码蛋白质的功能,可将 RNA 分为编码 RNA(coding RNA)和非编码 RNA(non-coding RNA,ncRNA),按照这样的划分,显然 mRNA 以外的所有 RNA,如 tRNA、rRNA、snRNA 和 miRNA 等都属于 ncRNA。而 ncRNA 还可以进一步分为管家 ncRNA(house-keeping ncRNA)和调控 ncRNA(regulatory ncRNA),前者呈组成型表达,是细胞的正常功能和生存所必需的,后者只在特定的细胞,或者在生物发育的某个阶段,或者在受到特定的外界刺激以后才表达,它们的表达能够在转录或翻译水平上影响到其他基因的表达。

e6-1 不同类型的 RNA 的功能和分布

Quiz1 e-1 中的 RNA 哪些属于管家 ncRNA,哪些属于调控 ncRNA?

第二节 核酸的一级结构

核酸的一级结构是指构成核酸的多聚核苷酸链上所有核苷酸或碱基的排列顺序。

图 6-1 为两种线形四聚核苷酸的结构。通过观察,可以发现多聚核苷酸链至少具有 3 个重要特征:①有两个不对称的末端,其中有一端的核苷酸 5′-OH 不参与形成 3′,5′-磷酸二酯键,此末端称为 5′端,另一端的核苷酸 3′-OH 不参与形成 3′,5′-磷酸二酯键,此末端称为 3′端。多聚核苷酸链的这种性质称为方向性或极性;②在生理 pH 下,多聚核苷酸链上的磷酸基团带有大量负电荷,因此核酸是一种多聚阴离子复合物;③每一条链上的核苷酸残基都有一定的排列顺序,这种排列顺序也就是核酸的一级结构。

5′端 A T G C 3′,5′-磷酸二酯键 3′端

四聚脱氧核苷酸

5′端 A U G C 3′,5′-磷酸二酯键 3′端

四聚核糖核苷酸

图 6-1 构成 DNA 和 RNA 核苷酸的结构和连接方式

除了线形核酸以外，自然界还有环形核酸，例如细菌的染色体 DNA、质粒 DNA（plasmid DNA）、叶绿体 DNA 和大多数线粒体 DNA 都属于环形 DNA。与线形核酸不同的是，环形核酸没有游离的 3′端和 5′端，或者说两个末端之间也形成了磷酸二酯键。

对于一个已知序列的核酸，可以用图 6-2 所示的方法来表示其一级结构：以垂直线代表核糖或脱氧核糖，斜线（中间有 P 或无）代表 3′，5′- 磷酸二酯键，短斜线代表两端。也可以直接用碱基的单字母缩写来表示，如果两端有磷酸基团，可用 p 表示；如果不确定是否有磷酸基团，可用“—”表示。按照惯例，应从 5′→ 3′书写核酸的一级结构。如果你一定要从 3′→ 5′书写，必须明确注明。有时，为了强调是 DNA 还是 RNA，可以在单字母缩写前加写 d 和 r。按照以上规则，图 6-1 左边的四聚脱氧核苷酸可写为 pApTpGpC 或 pATGC 或 pdApdTpdGpdC 或 pd（ATGC）；图 6-1 右边的四聚核糖核苷酸可写为 pApUpGpC 或 pAUGC 或 prAprUprGprC 或 pr（AUGC）。如果只对核酸的碱基序列感兴趣，那么可以直接使用碱基的单字母缩写表示一种核酸的一级结构。

Quiz2 GC 与 CG 序列是一样的吗？为什么？

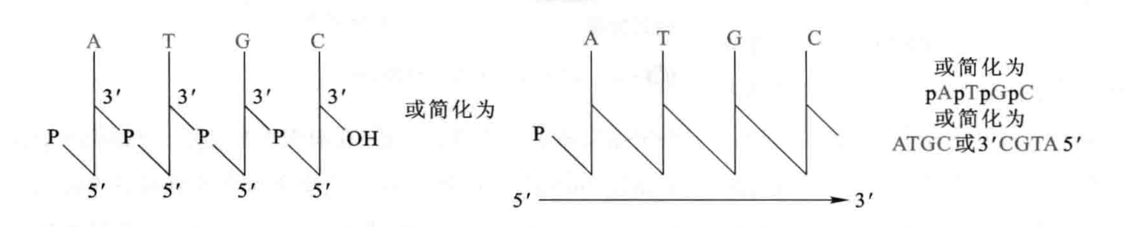


图 6-2 核酸一级结构的缩写

核酸一级结构的意义在于，生物体的遗传信息是储存在由 4 种核苷酸编码的特定的序列之中，而与高级结构无关。

第三节 核酸的高级结构

核酸的高级结构包括二级结构、三级结构，与蛋白质不同的是一般不认为核酸具有四级结构。

一、DNA 的高级结构

（一）DNA 的二级结构

DNA 的二级结构主要是各种形式的螺旋，特别是 B 型双螺旋，此外还有 A 型双螺旋、Z 型双螺旋、三螺旋和四链结构等。其中 B 型双螺旋的主要内容如下：

（1）DNA 由两条反平行的多聚脱氧核苷酸链组成，两条链相互缠绕，形成右手双螺旋。

（2）组成右手双螺旋的两条链在碱基序列上是互补的（complementary），它们通过特殊的碱基对结合在一起。碱基配对规则是一条链上的 A 总是与另一条链的 T，且 G 总是与 C 以氢键配对。其中 AT 碱基对有 2 个氢键，GC 碱基对有 3 个氢键（图 6-4）。这种配对方式被称作 Watson-Crick 碱基对。

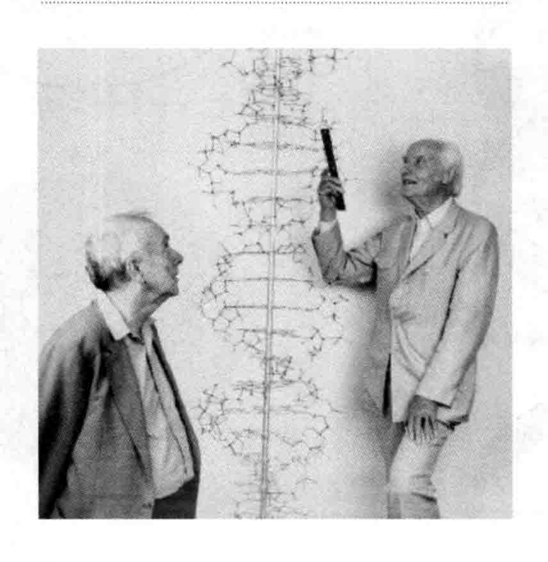


图 6-3 已步入古稀之年的 Watson（左）和 Crick（右）在讨论 DNA 双螺旋结构模型

有两个因素决定 Watson-Crick 碱基对的特异性，一是 A 和 T、G 和 C 之间在几何形状上的完美互补；二是它们之间能形成“绝配”的氢键。前一种因素可能更为

Quiz3 游离的碱基在溶液中可以随意通过氢键配对吗?

重要,因为有人用几何形状与T极为相似但缺乏氢键受体和供体的化合物——2,4-二氟甲苯,来代替双链DNA在少数位置上的T,结果双螺旋结构照样能够存在。

(3) 碱基对位于双螺旋的内部,并垂直于螺旋轴。碱基对之间通过疏水键和范德华力相互垛叠在一起,对双螺旋的稳定起重要作用。

(4) 双螺旋的表面含有明显的大沟(major groove)和小沟(minor groove),宽度分别为2.2 nm和1.2 nm(图6-5)。双螺旋上出现大、小沟的原因是,与每一个碱基对有关的两个糖苷键没有完全处于相对的位置。如果把螺旋看成一个圆筒的话,在对螺旋横切后,看到的两个糖苷键之间的夹角不是180°,而是120°的小角或240°的大角(图6-6)。当越来越多的碱基对垛叠在一起的时候,在碱基对一侧的小角形成小沟,在另外一侧的大角则形成大沟。显然,如果一个碱基对上的两个糖苷键夹角为180°的话,两个沟的大小和形状应该是一样的。

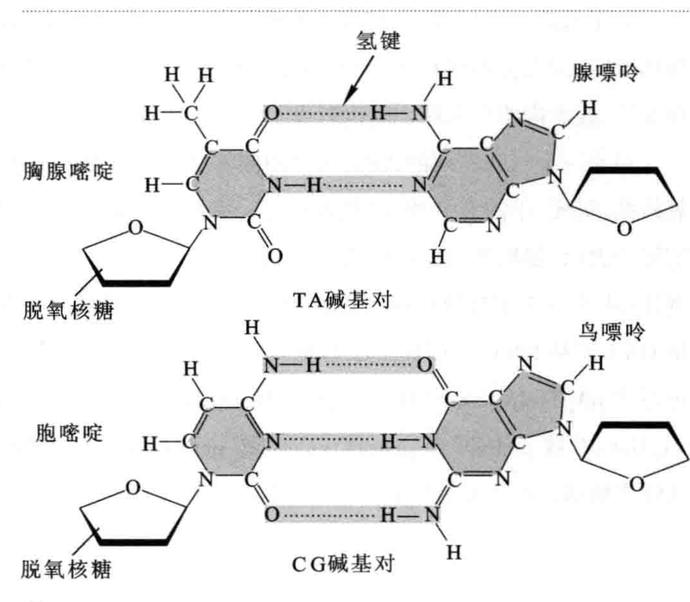


图6-4 AT和GC碱基对的配对性质

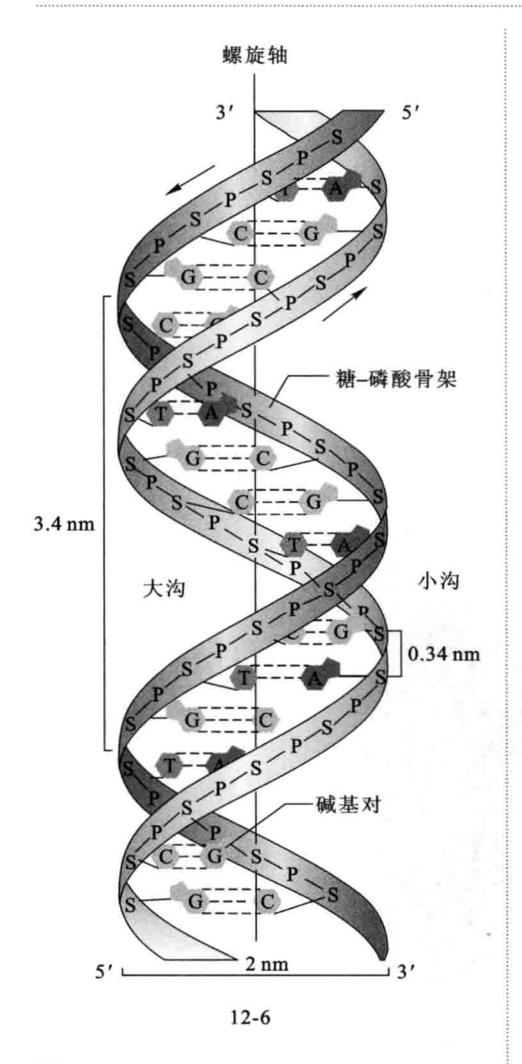


图6-5 B型DNA双螺旋结构的主要特征

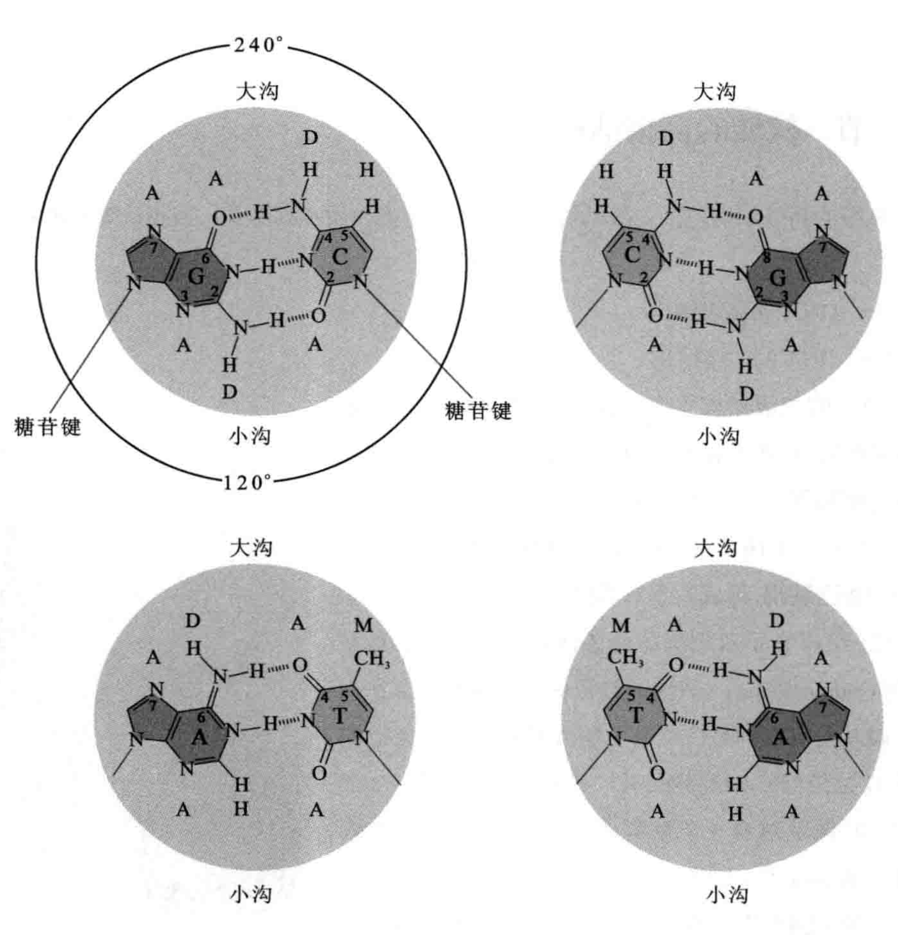


图6-6 DNA双螺旋大、小沟形成的原因以及里面展示的化学信息(Watson等,2007)

(5) 双螺旋的其他参数包括：相邻碱基对距离为 0.33 nm，相差约 36°。螺旋的直径为 2 nm，螺距为 3.32 nm，每一圈完整的螺旋含有 10 个碱基对（base pair，bp），然而细胞内的 DNA 双螺旋每一圈实际上含有 10.5 bp。

2. A 型双螺旋和 C 型双螺旋

B 型双螺旋仅仅是 DNA 双螺旋多种构象中的一种，在一定的条件下，双链 DNA 可以从 B 型转变成其他构象，例如 A、C、D、E、T 和 Z 等形式，但在正常的细胞环境中，能够存在的双螺旋只有 B 型、A 型和 Z 型（图 6–7），其中 B 型是细胞内最主要的形式，A 型一般与 RNA 有关系。

引起 DNA 双链构象改变的主要因素包括：①相对湿度；②盐的种类和浓度；③碱基组成和序列；④超螺旋的数量和方向。

框 6–2　生化探究——"大沟"和"小沟"中的秘密

双螺旋上的大沟和小沟是 DNA 分子对外展示信息和交流信息的窗口。大沟里含有更多的化学信息，因为在其中的每个碱基对能提供更多的氢键供体和受体以及疏水基团，这些信息可以被 DNA 序列特异性结合蛋白用来识别和结合特异性碱基序列。例如（图 6–6），AT 碱基对在大沟里展示的氢键供体（D）有与腺嘌呤 C6 相连的环外氨基，氢键受体（A）有腺嘌呤的 N7、胸腺嘧啶 C4 上的羰基 O，提供的疏水基团是与胸腺嘧啶 C5 相连的甲基；GC 碱基对在大沟里展示的氢键供体有与胞嘧啶 C4 相连的环外氨基，氢键受体（A）有鸟嘌呤的 N7、鸟嘌呤 C6 上的羰基 O，提供的疏水基团则是与胸腺嘧啶 C5 相连的 H。因此，在大沟里，不仅可以很好地区分 AT 和 GC，还可以区分 AT 和 TA、GC 和 CG。在这里，如果把特定的氢键受体（A）、氢键供体（D）、疏水基团甲基（M）和疏水基团氢原子（H）的组合看成是一种识别码，那么，AT 碱基对和 TA 碱基对的识别码分别是 ADAM 和 MADA，GC 碱基对和 CG 碱基对的识别码分别是 AADH 和 HDAA。这种识别码是非常重要的，因为一种 DNA 序列特异性结合蛋白不需要破坏 DNA 的双螺旋结构，就可以有效地识别特定的碱基序列。

与大沟相比，小沟里面展示出的化学信息要少，AT 碱基对能提供的氢键受体是腺嘌呤的 N_3、与胸腺嘧啶 C2 相连的羰基 O，无氢键供体，疏水基团是与腺嘌呤 N2 相连的 H；GC 碱基对能提供的氢键供体是与鸟嘌呤 C2 相连的环外氨基，氢键受体是鸟嘌呤的 N3 以及与胞嘧啶 C2 相连的羰基 O，无疏水基团。由此看来，在小沟里面区分不了 AT 和 TA 以及 GC 和 CG，因为 AT 和 TA 的识别码都是 AHA，GC 和 CG 的识别码都是 ADA。同时，小沟窄而浅，难以容纳氨基酸的侧链基团。正因为如此，DNA 序列特异性结合蛋白绝大多数是在大沟里面识别和结合 DNA 的，只有少数（如 TATA 盒结合蛋白）在小沟里识别碱基序列。但这些在小沟中起作用的蛋白质在识别的过程中，会撕开双链，以获取更多的信息。

促进 A 型 DNA 双螺旋（A–DNA）形成的主要因素是相对湿度的降低和盐浓度的上升。DNA 钠盐在相对脱水的条件下（低于 75% 相对湿度），可形成 A 型双螺旋。与细而长的 B–DNA 相比，A–DNA 更宽、更平、更短：每一圈螺旋为 11 bp，而不是 10 bp；螺距为 2.46 nm，而不是 3.32 nm；螺旋的直径为 2.6 nm，而不是 2.0 nm；碱基对不再近乎垂直于螺旋轴，而是相对于螺旋轴倾斜了 19°；相邻的碱基对距离 0.23 nm，而不是 0.33 nm；大沟窄而深，小沟宽而浅。

因为活细胞内充满着水，所以 A–DNA 基本上是不太可能存在的。但某些细菌孢子内的 DNA 是 A 型，这显然与孢子内部水分缺乏有关。除此之外，在 DNA 复制的时候，与 DNA 聚合酶活性中心结合的大概在 3 bp 内的双螺旋为 A 型。在基因转录的时候，DNA 与 RNA 形成的杂交双螺旋为 A 型。RNA 双螺旋也为 A 型。双链 RNA 或者 DNA–RNA 杂交双链之所以以 A 型双螺旋存在，是因为 RNA 的 2′–OH 造成的空间位阻阻止了 B 型双螺旋的形成。

盐的种类能够改变 DNA 双螺旋构象，例如，DNA 锂盐在 66% 的相对湿度条件下，可形成 C 型双

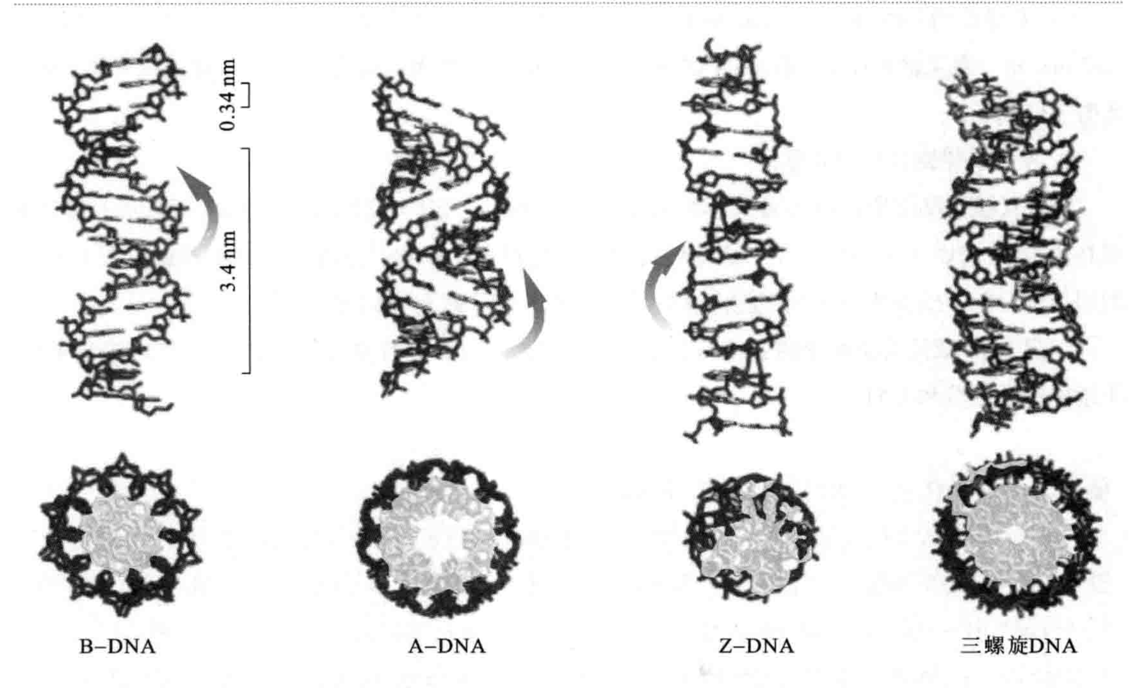


图 6-7 A B Z 三种双螺旋及三螺旋 DNA 的棍式模型(Watson 等,2007)

螺旋。但这种 DNA 没有任何生物学的意义,因为细胞内的 DNA 不可能处在这样的条件下。

3. Z 型双螺旋

Z-DNA 最早是在体外得到的一种双螺旋。1972 年,Fritz M. Pohl 等人发现,人工合成的由 GC 相间排列的多聚核苷酸(GCGCGC)在高盐的条件下,旋光性会发生改变。MIT 的 Alexander Rich 随后对上述六聚脱氧核苷酸的单晶进行了 X 射线衍射分析,提出了 Z-DNA 模型。其结构特点包括:

(1) 戊糖磷酸骨架呈锯齿状或“Z”字形(zigzag) 其中的核糖和 C 以反式连接,和 G 以顺式连接。

(2) 左手螺旋 G 的糖苷键呈顺式,不仅使螺旋旋转方向发生了改变,还使 G 残基位于分子表面。

(3) 分子外形呈波形 这是由于碱基对向螺旋的外表面前移,中轴不再像 B-DNA 位于碱基对之间,而是移向了小沟。

(4) 小沟变深变窄 脱氧胞苷酸的糖环 C2′为内式,碱基为反式,使糖环转离小沟,而脱氧鸟苷酸的糖环 C3′为内式,碱基为顺式,使糖环弯向小沟。胞嘧啶的 C5 和鸟嘌呤的 N7、C8 原子填满了大沟,并指向表面,使大沟变得不明显,而小沟变得非常深但较窄。

(5) 螺旋的各种参数发生了变化(表 6-2)。

Z-DNA 在体外的存在需要满足两个条件:

(1) 盐浓度高或者乙醇存在时 如果是 NaCl,其浓度需超过 2 mol/L;如果是 $MgCl_2$,其浓度要超过 0.7 mol/L。Z-DNA 的形成导致带负电荷的磷酸根距离更近了(Z-DNA 为 0.8 nm,B-DNA 为 1.2 nm),而高的盐浓度能够减少两条链之间靠得最近的磷酸基团间的排斥力。

(2) 嘌呤 - 嘧啶相间排列,例如多聚(dG-dC) 目前认为在适当的盐浓度下,任何不少于 6 bp 的嘌呤 - 嘧啶交替排列顺序都能形成 Z-DNA。有时并不严格的嘧啶 - 嘌呤相间序列,例如含有 2 个 5 - 甲基胞嘧啶(5-methylcytosines,m^5C)的六聚核苷酸 m^5CGATm^5CG,在生理盐水的浓度下也能形成 Z-DNA。事实上,体内 m^5C 上的甲基被认为有助于 B 型向 Z 型的转变,这是因为在 B-DNA 上,疏水的甲基伸入到大沟内水溶性环境之中,从而不利于 B-DNA 的稳定。而在 Z-DNA 中,同样的甲基却能形成一种稳定性的疏水“补丁”(hydrophobic patch)。

生命科学的研究思路总是这样,当科学家在体外得到某种结构的时候,首先想到的是体内是否存在类似的或同样的结构,如果在体内发现存在这种结构,又想知道它在体内的功能。那么体内有没有

表 6-2 A B Z 三种双螺旋的比较

参数	A 型双螺旋	B 型双螺旋	Z 型双螺旋
外形	短而宽	长而瘦	长而细
每 bp 上升的距离	0.23 nm	(0.332 ± 0.19) nm	0.38 nm
螺旋直径	2.55 nm	2.37 nm	1.84 nm
螺旋方向	右手	右手	左手
螺旋内每重复单位的 bp 数	1	1	2
每圈 bp 数	≈11	≈10	12
碱基夹角	32.7°	34.6°	60°/2
螺距	2.46 nm	3.32 nm	4.56 nm
碱基对倾角	19°	1.2° ± 4.1°	– 9°
螺旋轴位置	大沟	穿过碱基对	小沟
大沟	极度窄，很深	很宽，深度中等	平坦
小沟	很宽，浅	窄，深度中等	极度窄，很深
糖苷键构象	反式	反式	C 为反式，G 为顺式
糖环折叠	C3′内式	C2′内式	嘧啶 C2′内式，嘌呤 C3′内式
存在条件	双链 RNA，RNA-DNA 杂交双链，低湿度 DNA(75%)	双链 DNA(高湿度，92%)	嘧啶和嘌呤交替存在的双链 DNA 或 DNA 链上嘧啶和嘌呤交替存在的区域

天然的 Z-DNA 存在？如果有，其功能又是什么？

Rich 使用荧光标记的 Z-DNA 特异性抗体，发现在某些生物体内的确存在 Z-DNA。例如，果蝇的 X 染色体上就存在 Z-DNA，而人类基因组上类似的可暂时形成 Z-DNA 的片段约有 10^5 bp，但细胞内 Z-DNA 形成的条件与体外有所不同。例如，体内的盐浓度很难达到非常高的水平，而体内带正电荷的多胺化合物，如精胺(spermine)和亚精胺(spermidine)的存在，一样可与磷酸基团结合，减少负电荷的排斥作用，使 B-DNA 转变成 Z-DNA。此外，某些蛋白质(如 Z-DNA 结合蛋白)带有正电荷，与 DNA 结合可在 DNA 周围创造高盐浓度的微环境，这也是在活细胞中形成 Z-DNA 的原因之一。有时体内负超螺旋的存在也能促进 Z-DNA 的形成，这是因为负超螺旋引起的扭曲张力可稳定 Z-DNA 的存在。抗体结合实验表明，将多聚(dG-dC)插入到质粒中，质粒在松弛状态下不能结合 Z-DNA 抗体，但如果处于负超螺旋状态，则能结合。

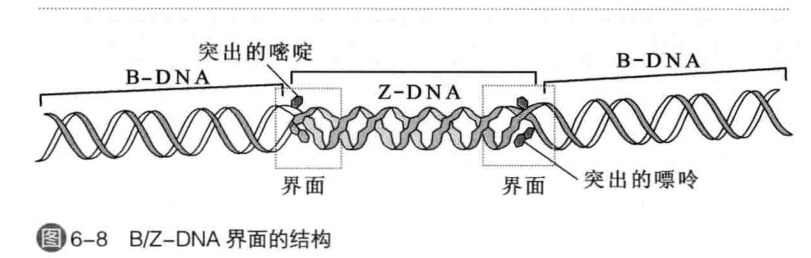


图 6-8 B/Z-DNA 界面的结构

如果细胞内的 DNA 在某段序列形成 Z-DNA，那这段 Z-DNA 在两端与周围 B-DNA 的交界处就会各有一对碱基从内部被挤出来(图 6-8)，碱基的突出与磷酸核糖骨架在两种双螺旋过渡的时候发生弯曲和回转有关。

关于体内 Z-DNA 形成的生物学意义，现在一般认为与基因的表达调控有关。已发现有利于 Z-DNA 形成的碱基序列经常位于控制基因转录起始的启动子区域，而该区域的 DNA 在基因转录或染色质重塑过程中易形成负超螺旋。也许 DNA 的局部区域发生 Z-DNA 和 B-DNA 的可逆性转变可充当控制基因表达的开关。当一个 DNA 分子的某个区域从 B-DNA 转变成 Z-DNA 以后，细胞内的某些调节蛋白特异性地与 Z-DNA 结合，从而调控基因表达。例如，猿猴肾病毒(SV40) DNA 上能够增强基因表达的增强子序列中就有此结构。研究发现，真核细胞内 DNA 上有许多 m^5C，既然 m^5C 有利于 Z 型双螺旋的形成，而甲基化能够调控基因的表达，这就说明 Z-DNA 与基因表达之间肯定存在相关性。

Z-DNA 除了参与基因表达调控以外，还可能通过释放来自螺旋本身的拓扑学张力而促进同源

重组。

4. 支持DNA双螺旋结构的证据

Watson 和 Crick 提出的 DNA 双螺旋结构模型主要建立在 X 射线衍射数据和 Chargaff 规则的基础之上。

(1) X射线衍射数据　从 1950 年到 1953 年，Rosalind Franklin 和 Maurice Wilkins 使用 DNA 湿纤维(类似于晶体)进行了 X 射线衍射分析，所取得的一些数据表明，DNA 具有简单、有规律的重复结构单元，详细分析见图 6-9。

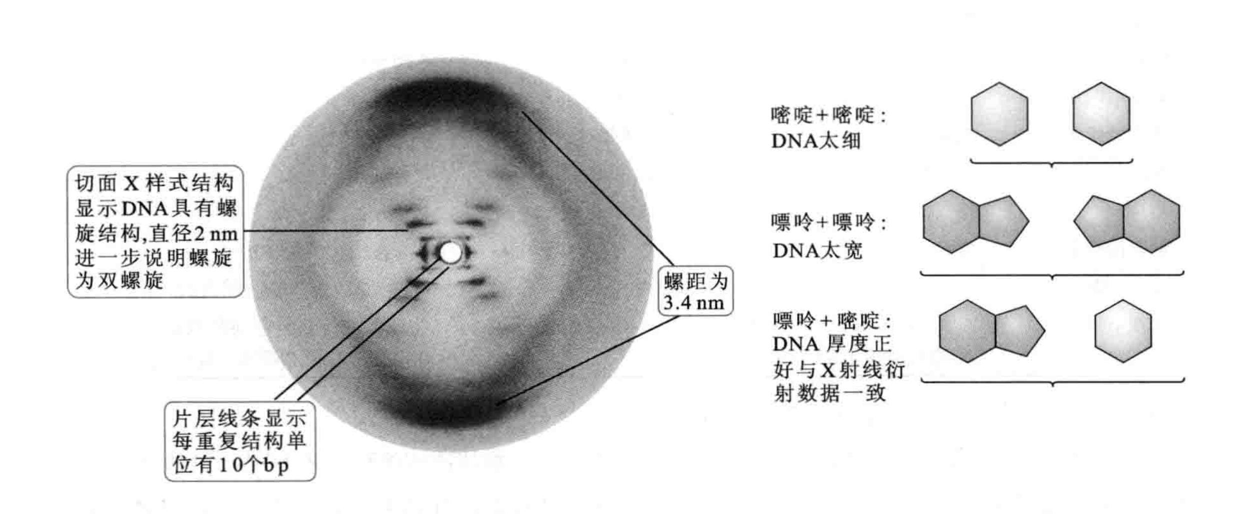


图6-9　DNA 湿纤维的 X 射线衍射图及其解释

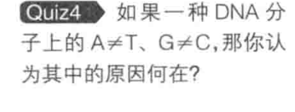
(2) Chargaff 规则　1949 年，Erwin Chargaff 应用纸层析及紫外分光光度计，对不同生物来源的 DNA 的碱基组成进行了定量测定，结果发现在所测定的每一种生物的 DNA 分子上，A 和 G 的含量都分别等于 T 和 C，且 A+G=C+T(表 6-3)。此外，DNA 的碱基组成具有物种特异性，但无组织和器官的特异性，并且年龄、营养状况和环境等因素不影响 DNA 的碱基组成。

Quiz4　如果一种 DNA 分子上的 A≠T、G≠C，那你认为其中的原因何在?

表 6-3　不同物种 DNA 的 A : T、G : C 和 A : G 的比值

物种	A : T	G : C	A : G
人	1.00	1.00	1.56
大麻哈鱼	1.02	1.02	1.43
小麦	1.00	0.97	1.22
酵母	1.03	1.02	1.67
大肠杆菌	1.09	0.99	1.05
黏质沙雷菌	0.95	0.86	0.70

上述 DNA 分子中四种碱基组成表现出来的规律称为 Chargaff 规则(Chargaff's rule)。尽管 Chargaff 对这一发现谨慎地猜测为“这是 DNA 结构某些重要特征的反映”，然而他本人并没有揭示出其中的生物学意义。Watson 和 Crick 则正好利用此规则来说明 DNA 双螺旋结构中 A 与 T、G 与 C 配对的特征。



罕见的亚胺基异构体

图6-10　错误配对的 CA 碱基对

(3) 碱基的互变异构　碱基配对的特异性在很大程度上是由碱基之间的氢键决定的，但碱基具有酮式和烯醇式或氨基式和亚氨基式(参看第五章“核苷酸”)，两种形式的碱基在氢键的供体和受体上有所变化。GC 和 AT 配对是碱基在酮式结构下的配对，这也是 DNA 分子正确的配对方式，因为体内的碱基主要形式正是酮式。但如果碱基变成烯醇式，就会发生其他形式的配对，如 CA 配对(图 6-10)。事实证明，在 DNA 复制过程中，作为模板的某些碱基“不巧”处于烯醇式或亚氨基式，结果导致复制

出现错配(参看第三十三章“DNA 复制”)。这说明在 DNA 分子中,碱基的确是以配对的形式存在的,这与 Chargaff 规则也是一致的。

5. 稳定双螺旋结构的因素

DNA 的双螺旋结构通常是很稳定的,起稳定作用的因素包括:

(1) 氢键,包括螺旋内部的氢键和螺旋外部的氢键　螺旋内部的氢键是碱基对之间的氢键,外部氢键在戊糖 – 磷酸骨架上的亲水基团与周围的水分子之间形成。在双螺旋中,嘧啶和嘌呤之间的距离正好与一般氢键的键长(0.27 nm)差不多,且氢键供体原子和氢键受体原子处于一条直线上,利于形成氢键。若嘌呤与嘌呤或嘧啶与嘧啶配对,其空间的大小都不适合形成氢键。

氢键固然重要,但它们主要决定碱基配对的特异性,而对双螺旋稳定性的贡献不是最重要的。对双螺旋稳定性起决定性作用的是碱基堆积力。

(2) 碱基堆积(base-stacking)力　这是碱基对之间在垂直方向上的相互作用所产生的力,包括疏水作用和范德华力。

虽然多聚核苷酸链的磷酸核糖骨架是高度亲水的,但碱基杂环 π– 电子云导致嘧啶和嘌呤本身具有一定程度的疏水性。在水溶液中,这些疏水基团会自发聚集在一起。当疏水基团因聚集而相互靠近到一定的距离时,自然就会产生范德华力,这与球状蛋白质形成疏水核心的机制一样。从热力学的角度来看,DNA 形成双螺旋能使高度亲水的磷酸基团与水的接触增加到最大限度,同时也使碱基与水的接触减少到最小限度。

在双螺旋中,相邻碱基对的间隔是 0.34 nm,而范德华力的平均半径(指引力和斥力正好相平衡的距离)为 0.17 nm,因此相邻碱基对会产生范德华引力。

碱基间相互作用的强度与相邻碱基之间环重叠的面积成正比,其总的趋势是嘌呤与嘌呤之间 > 嘌呤与嘧啶之间 > 嘧啶与嘧啶之间。另外,碱基的甲基化也能提高碱基堆积力。

(3) 磷酸核糖骨架上负电荷的中和　在生理 pH 下,DNA 主链上的磷酸基团带有大量的负电荷,因此两条链之间存在着静电的排斥作用。溶液中的阳离子(如 Mg^{2+} 或多胺类化合物)可以中和磷酸基团所带的负电荷,消除静电斥力。当离子浓度降低,中和作用减弱时,排斥力会加大,DNA 双螺旋稳定性就降低了。

在上述几种因素的作用下,一段由 20 bp 组成的双螺旋在室温下已相当稳定。正因为如此,在设计聚合酶链式反应的引物的时候,要让引物与 DNA 模板形成稳定的双螺旋结构,引物长度至少是 16 个碱基,但为了增加特异性,最好达到 20~24 个碱基。

6. DNA 双螺旋结构的意义

DNA 双螺旋结构发现的意义在于为生物学家揭示遗传物质的遗传、复制、修复、多样性以及物种的进化机制提供了重要的线索。从更广泛的意义来看,该发现将遗传学与生物化学、细胞生物学和生理学等学科结合到一起,并直接导致分子生物学的诞生及之后的迅速发展。简言之,在 DNA 双螺旋结构模型中,最重要的内容应该是互补的概念。互补结构对于理解 DNA 的复制、重组、修复、转录、转录后加工和翻译等机制都很重要。

可以说,在发现 DNA 双螺旋结构之前,想推测出 DNA 作为遗传物质如何能够行使它的生物学功能几乎是不可能的。但在发现 DNA 由两条链组成,并且两条链的序列是互补的以后,一切就变得很容易了。一个 DNA 分子上的两条链都可以分别作为模板,即新链的序列按照碱基互补配对原则由模板链决定。结果就是,DNA 链上的任何一部分序列都可以用来作为复制新序列的模板,或者去识别另外一个 DNA 分子上的互补序列后再配对结合。这两点对于 DNA 复制和重组至关重要(参看第三十三章“DNA 复制”和第三十五章“DNA 重组”)。此外,DNA 双螺旋结构对于 DNA 损伤的修复机制也同样重要。当 DNA 受到损伤以后,机体内修复系统通常会把受到损伤的序列切除,然后以其正常的互补序列为模板,重新合成出正确的序列。

7. DNA 的非标准二级结构

除了上述 3 种最常见的标准二级结构以外，细胞内的 DNA 在特殊的条件下还可能形成其他几种非标准的二级结构，如弯曲（bending）、十字形（cruciforms）、三螺旋（triple helix）、滑动（slipped mispaired DNA，SMP-DNA）错配和碱基翻转（base flipping）等。这些特殊的条件包括：DNA 受到某些蛋白质的作用（如组蛋白）；DNA 本身具有几种特殊序列模体（图 6-11），例如反向重复（inverted repeat）、回文序列（palindromic sequence）、镜像重复（mirror repeat）、直接重复（direct repeat）、高嘌呤序列、高嘧啶序列、富含 A 序列和富含 G 序列。

反向重复

CGAACGTCCTATTGGGGACGTTCC
GCTTGCAGGATAACCCCTGCAAGG

回文序列

AACGAATTCCCC
TTGCTTAAGGGG

镜像重复

AGGGGA
TCCCCT

直接重复

CGAACGTCC TATTGG GAACGTCCC
GCTTGCAGG ATAACCCTTGCAGGG

图 6-11　四种特殊的碱基序列

（1）弯曲　当 DNA 的一条链在某区段含有成串的 A 序列（4~6 个 A），且相邻的串间隔 10 个碱基的时候，该区段 DNA 很容易形成弯曲。有时 DNA 受到某些蛋白质的作用也会形成弯曲。例如，真核生物的 TATA 盒结合蛋白可导致 DNA 在启动子区域发生弯曲，大肠杆菌 cAMP 受体结合蛋白可导致 DNA 在乳糖操纵子处发生弯曲。已有证据表明，在 DNA 复制、转录和定向重组过程中，DNA 在局部形成弯曲可能是蛋白质与 DNA 序列之间相互作用的要素之一。此外，DNA 发生碱基错配或 DNA 受到紫外线的照射发生损伤时也可以形成弯曲，此时形成的弯曲实际上是作为体内 DNA 修复系统识别损伤的信号。

（2）十字形　如图 6-12 所示，DNA 内部一些含有反向重复的区域在解链以后，可通过链内碱基互补配对形成一种十字形的二级结构。由于十字形结构在两端含有 6~7 个没有配对的碱基，因此在热力学上并不是一种稳定的结构。

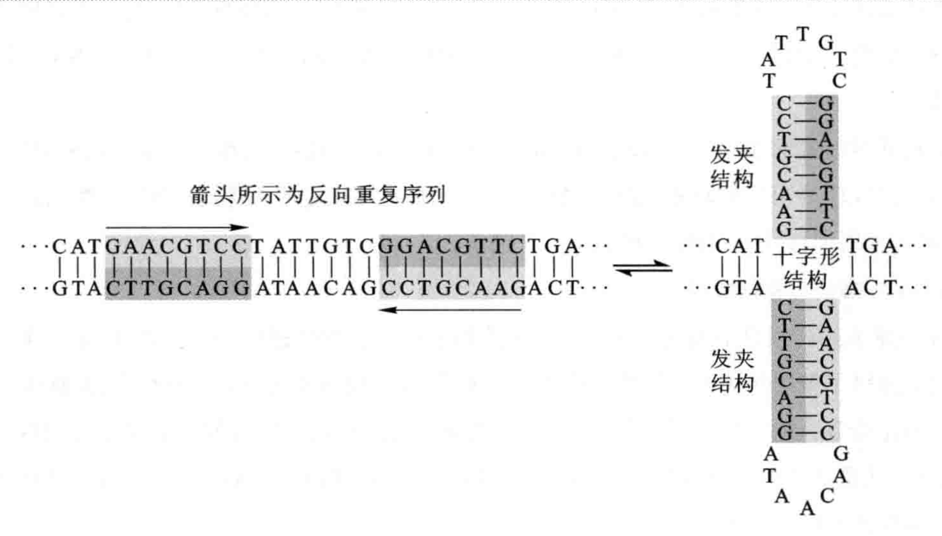


图 6-12　十字形 DNA 的形成

在人类基因组 DNA 的复制起始区和转录调控区，有许多反向重复序列，因此，这些区域有可能形成十字形结构，其功能也许是充当控制 DNA 复制和基因转录的开关。

（3）三螺旋与 H-DNA　早在 20 世纪 50 年代和 60 年代，已有人发现，将某些特殊序列的多聚脱氧核苷酸链按照一定的比例混合在一起，可以形成三螺旋结构即铰链 DNA（hinged DNA，H-DNA），例如将多聚 dA 和多聚 dT 按照 1∶2 的浓度混合在一起可形成 ATT 三螺旋。

进一步研究表明，三螺旋的形成需要至少 DNA 的一条链全部是由嘌呤碱基组成。实际上，三螺旋是在双螺旋的基础上形成的：位于双螺旋大沟之中的嘌呤碱基具有再形成两个氢键的潜在位点（称

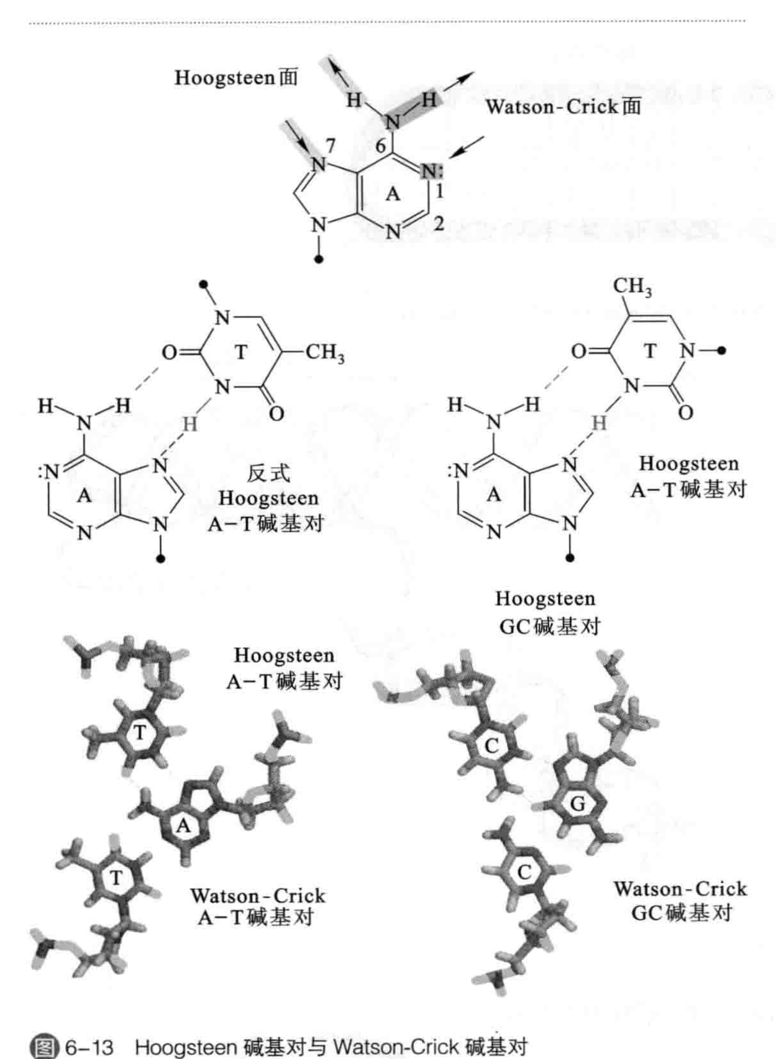


图 6–13　Hoogsteen 碱基对与 Watson-Crick 碱基对

为 Hoogsteen 面)。若是 G,两个位点为 N7 和 O6;若是 A,两个位点是 N7 和 6 号位的 NH_2,在双螺旋的一条链为多聚嘌呤核苷酸的时候,大沟中的嘌呤碱基就可以通过 Hoogsteen 碱基对(T 与 A 配对,质子化的 C 与 G 配对)与第三条链形成三螺旋。

Hoogsteen 碱基对有顺式和反式两种。如图 6–13 所示,反式是指第三条链与嘌呤链呈反平行排列,顺式是指第三条链与嘌呤链呈平行排列。

在三螺旋结构中,碱基堆积力对其稳定性有一定的贡献,但由于三条链之间存在更强的电荷排斥,稳定性没有双螺旋好。当然,当 DNA 处于高盐浓度下,其链上的负电荷多数被多价阳离子中和,形成三螺旋将会变得容易。此外,负超螺旋或低 pH 也有利于 H–DNA 的形成。如果双螺旋 DNA 内部含有由嘌呤构成的镜像重复序列时,三螺旋可以在一个 DNA 分子内形成(图 6–14)。

据估计,这种由嘌呤组成的镜像重复序列在人 DNA 上大概每 140~150 kb 出现一次,也就意味着 H–DNA 可能广泛地存在于基因组 DNA 中。在细胞内,H–DNA 经常出现在 DNA 复制、转录和重组的起始位点或调节位点,因此有人推测,第三股链的存在可能使一些转录因子或 RNA 聚合酶难以与该区段结合,从而阻遏有关基因的表达。原则上,人工合成的单链 DNA 如果能与体内的某段 DNA 序列形成三螺旋,就能阻断特定的基因表达,也就能治疗一些因特定基因过度表达引起的疾病。

(4) 碱基翻转　有时候,DNA 双螺旋上的某个碱基离开它的“配偶”,突出在双螺旋之外,这种现象称为碱基翻转(图 6–15)。一个碱基发生翻转的时候,会造成相邻碱基对的扭曲。碱基翻转对于细胞的某些功能是很重要的。例如,参与同源重组的酶需要通过碱基翻转寻找同源的序列,催化碱基修饰的酶需要碱基通过翻转落入它的活性中心被化学修饰,参与碱基切除修复的 DNA 糖苷酶需要受损伤的碱基通过翻转进入它的活性中心被切除。最后一种情况的实例就是 DNA 分子上出现的 U 是通过这种方式被尿嘧啶 –DNA 糖苷酶切除掉的(参看第三十四章“DNA 的损伤、修复和突变”)。

(5) 滑移错配 DNA　含有直接重复序列的 DNA 可以形成一种叫“滑移错配”的二级结构。形成这种结构的原因是该区段 DNA 先发生解链,在重新缔合的时候,一段重复单元内的核苷酸序列因滑移与另一段重复单元内的互补序列发生错配,从而形成两个环(图 6–16)。因滑移的方式不同,可形成两种 SMP–DNA。若体内的 DNA 形成上述结构,会导致某些基因发生移框突变(参看第三十四章“DNA 的损伤、修复和突变”)。

除了以上 5 种非标准的二级结构以外,DNA 在特殊的条件下,还可能形成一种四链结构

Quiz5　在 DNA 双螺旋结构提出之前,Pauling 与 Watson、Crick 都提过一种三螺旋结构。与这里的三螺旋结构不同的是,他们认为碱基在外、磷酸基团在螺旋内部。你认为这种三螺旋的结构为什么很难存在?

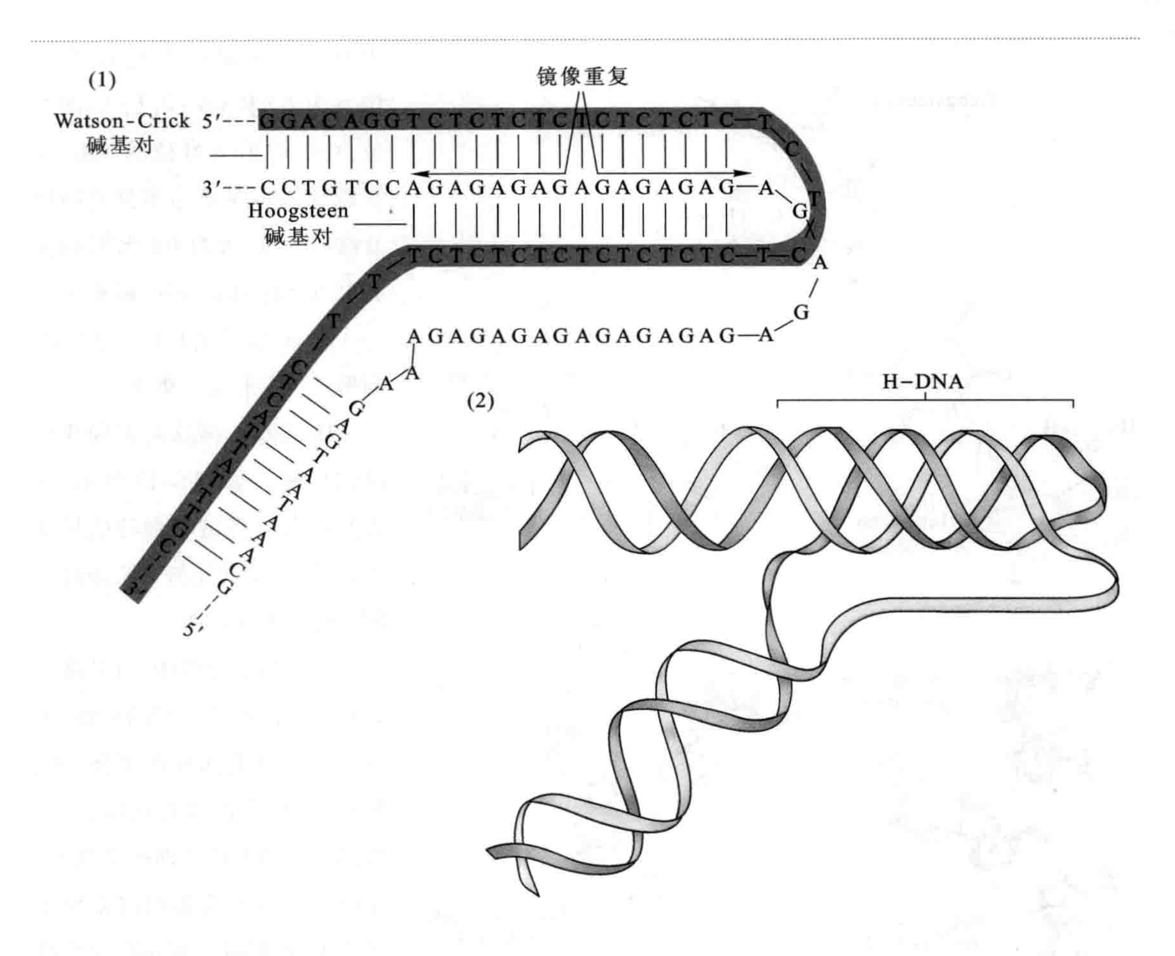


图 6-14　三螺旋 DNA(H-DNA)(Nelson 等,2008)

e6-2 DNA 的四链结构

(quadruplex)。这种四链结构也许存在于真核细胞内,并具有特殊的功能。

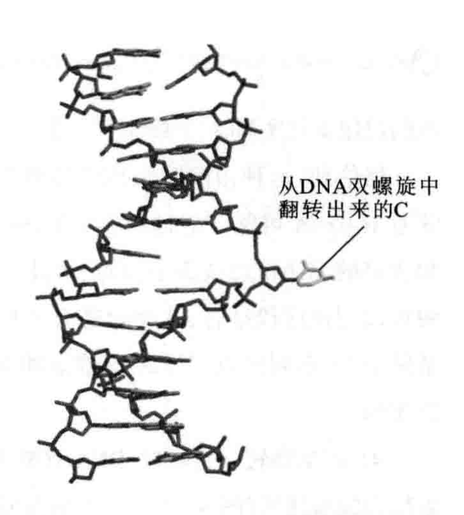


图 6-15　DNA 双螺旋上某一个位置的 C 发生的翻转

（二）DNA 的三级结构

如图 6-17 所示,DNA 可以以两种形式存在,右为松弛型(relaxed),左为超螺旋(supercoiling)。在松弛型状态下,DNA 以正常的 B 型双螺旋存在,每一圈 10 个碱基对,这时候双螺旋的能量状态最低。当通过某种手段使得 DNA 双螺旋每一圈的碱基对数目多于或少于 10 对,将导致 DNA 双螺旋缠绕过度或缠绕不足。如果这时 DNA 两端被固定或者 DNA 本来是共价闭环(covalently close circular DNA,cccDNA)的,DNA 将会因内部的张力无法释放而自发地形成超螺旋结构。

DNA 超螺旋分为正超螺旋和负超螺旋,前者为左手超螺旋,由 DNA 双螺旋过度缠绕引起;后者为右手超螺旋,由 DNA 双螺旋缠绕不足引起。

超螺旋 DNA 可以通过以下几种参数定量地表示(图 6-18)。

(1) 连环数(linking number,*L*)　*L* 是拓扑学中的一个重要术语,用在超螺旋 DNA 上是指双螺旋 DNA 的一条链环绕另一条链的次数。对于负超螺旋,*L* 为正值。对于一个共价闭环 DNA(covalently—closed circular,cccDNA)来说,其 *L* 值是不变的。例如,一个 5 400 bp 大小的松弛型 B 型 cccDNA,其 *L* 值应该是 5 400/10=540。

(2) 扭转数(twisting number,*T*)　*T* 是指 DNA 的双螺旋的总转数。对于负螺旋,*T* 为正值。对于一个倾向形成 B 型双螺旋的 DNA 来说,其最佳 *T* 值为碱基对数 /10。

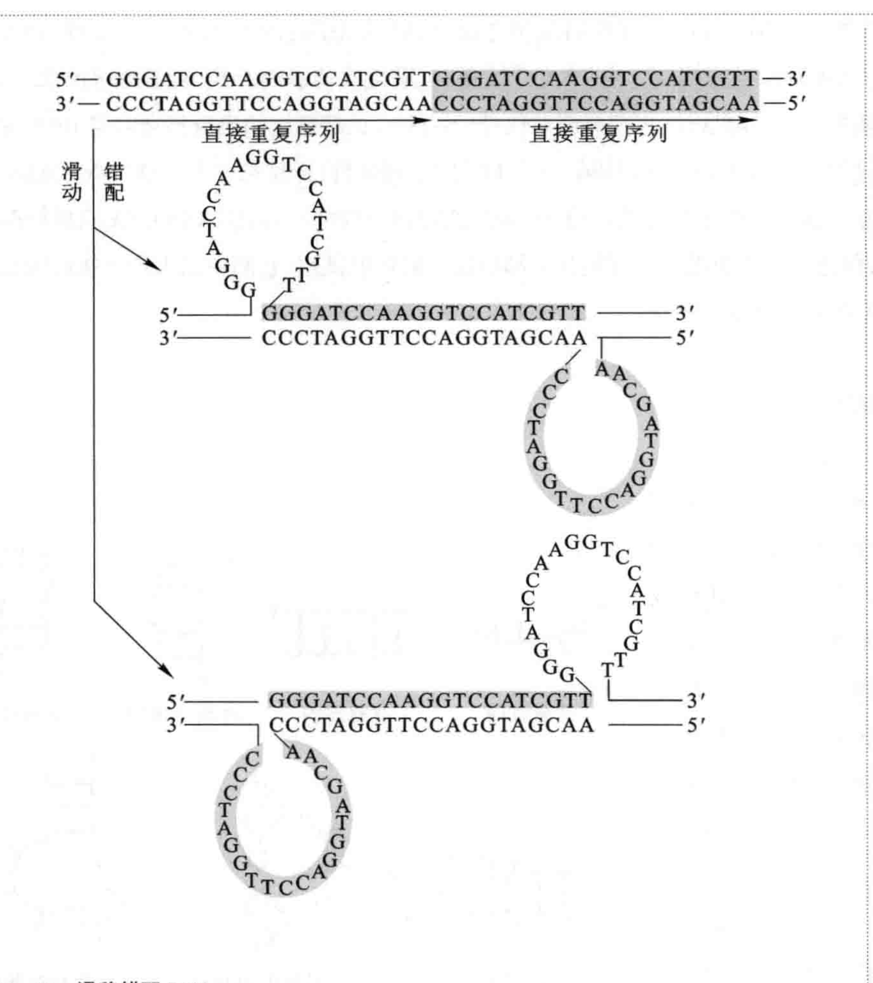
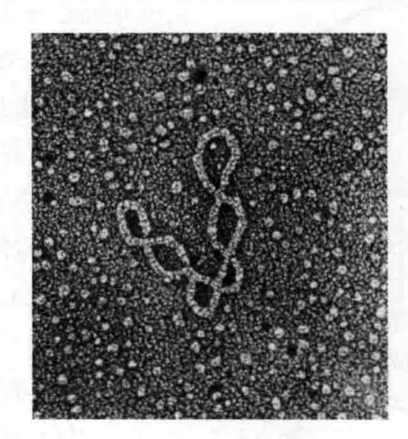


图 6-16 滑移错配 DNA


超螺旋DNA

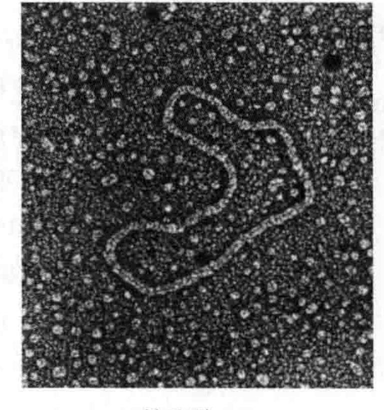
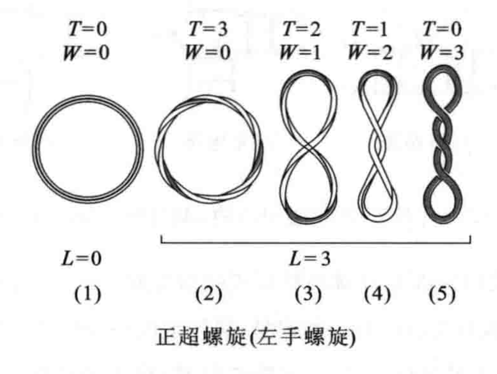

松弛型DNA

图 6-17 超螺旋 DNA 和松弛型 DNA



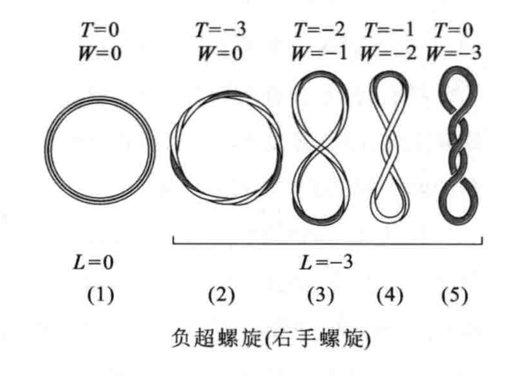


图 6-18 正超螺旋和负超螺旋

(3) 缠绕数(writhing number, W) 即超螺旋数,是指双螺旋的轴环绕超螺旋轴的次数。负超螺旋的 W 值为负数,正超螺旋的 W 值则为正数。

根据拓扑学理论,L、T 和 W 之间的关系可用以下方程表示:$L=T+W$。T 和 W 可以为小数值,但 L 必须为整数值。

(4) 比连环差(specific linking difference, λ) 该数据用来表示 DNA 的超螺旋密度(superhelix density),用 λ 表示。

$\lambda=(L-L_0)/L_0$,其中 L 是指某一超螺旋 DNA 的 L 值,L_0 是指松弛环形 DNA 的 L 值。

一般天然 DNA 分子中 $\lambda=-0.05=5\%$ 负超螺旋,例如 SV40,其基因组大小为 5 226 bp,$T=522$,$W=-26$($L=496$),$\lambda=-0.05$。

负超螺旋 DNA 是由两条链的缠绕不足引起,很容易解链,因此有利于 DNA 的复制、重组和转录。

Quiz6 你认为何种生物体内的 DNA 在没有复制之前极有可能以正超螺旋的形式存在?为什么?

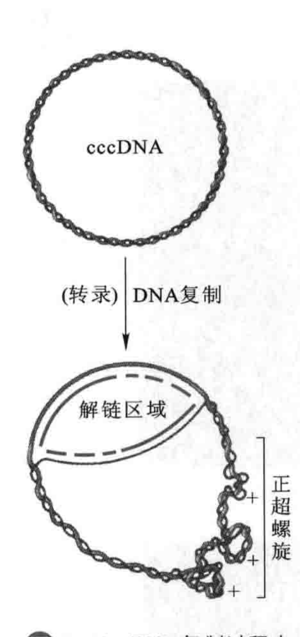


图 6-19 DNA 复制过程中正超螺旋 DNA 的形成

正因为如此，绝大多数生物体内的 DNA 在没有复制之前，均以负超螺旋的形式存在，以方便 DNA 在复制、重组或转录的启动阶段进行解链。当 DNA 开始复制、重组或转录的时候，随着解链的深入，原来的负超螺旋会逐渐被消耗，并最终被正超螺旋取代（图 6-19）。正超螺旋的出现将会阻碍 DNA 继续复制和转录，幸好细胞内存在 DNA 拓扑异构酶，可及时清除正超螺旋（详见第三十三章“DNA 复制”）。

某些与碱基对差不多大小的多环芳香族分子（如溴乙锭和吖啶橙）可以插入到 DNA 双螺旋两个相邻的碱基对之间，促进产生正超螺旋。这些分子都是强烈的致癌试剂，它们在插入到双螺旋内部以后，很容易诱发 DNA 在复制时发生突变。

二、RNA 的高级结构

（一）RNA 的二级结构

RNA 的二级结构主要取决于其碱基组成，即其一级结构。在知道一种 RNA 的一级结构以后，可以对其形成的二级结构进行很好地预测，目前有许多专业的软件可以进行这样的工作，如 RNA Draw 和 RNA Structure，也有专门的网站免费提供这样的服务，如 http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi。

RNA 二级结构的多样性可以和蛋白质相提并论。少数病毒的基因组 RNA 由两条互补的链组成 A 型双螺旋。多数 RNA 仅由一条链组成，它们的二级结构主要是由链内碱基的互补性决定的：互补的碱基之间可以配对形成链内 A 型双螺旋，非互补的碱基则游离在双螺旋之外，以凸起或环的形式存在，形成多种形式的二级结构（图 6-20）。

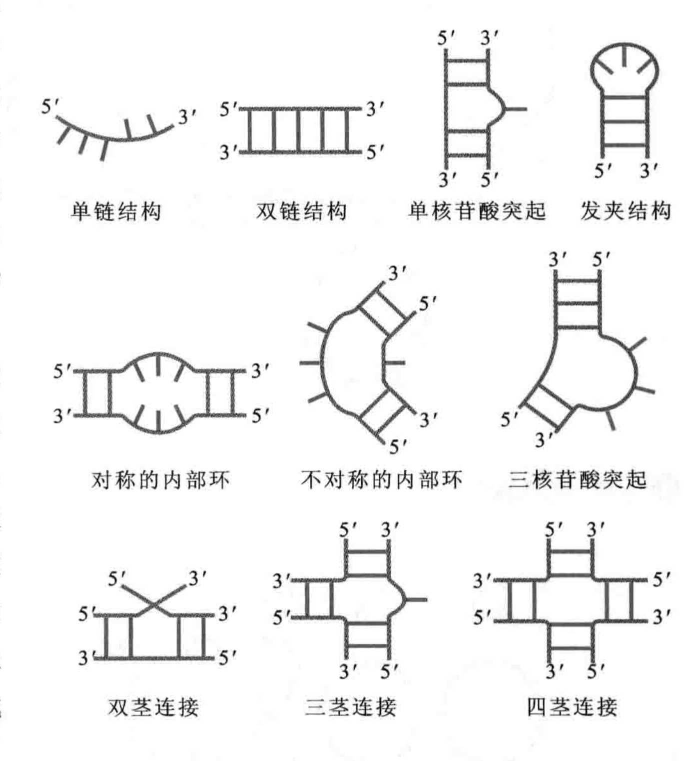


图 6-20 十种不同形式的 RNA 的二级结构（Alberts 等，2007）

Quiz7 RNA 分子中的哪一种二级结构与蛋白质中的 β 突起相似？

在这些二级结构中，发夹结构最为常见。发夹结构包括两个部分，一部分是由一段由标准互补碱基对形成的双螺旋，另一部分是双螺旋两股互补序列之间的一段由 3~5 个没有配对的碱基组成的环。最常见、最稳定的环是由 4 个碱基——UNCG、GNRA 或 CUYG（N 表示任何碱基，R 代表嘌呤碱基，Y 代表嘧啶碱基）构成的四环（tetraloop）。以 GNRA（如 GAGA、GCAA 和 GAAA）为例，环上的第一个碱基（G）和第四个碱基（A）形成不同寻常的 GA 碱基对，此外，在 G 和磷酸基团之间的一个氢键、R 与 2′- 核糖羟基的氢键以及碱基间的堆积对四环结构均有稳定作用。

图 6-21 GNRA 四环结构

在 RNA 双螺旋中，常常可以发现 GU 碱基对（图 6-22），这为单链 RNA 形成链内双螺旋创造了更多的机会。与 GC 碱基对不同的是，GU 碱基对只有 2 个氢键。不同种类的 RNA 的全部二级结构实际就是上述各种二级结构不同形式的组合。

Quiz8 为什么 RNA 结合蛋白通常在小沟中与 RNA 发生作用？

下面以细胞内 3 种最常见的 RNA，以及 DNA 在转录时生成的 RNA 产物与 DNA 模板形成的杂交双链为例，对于 RNA 的二级结构做进一步的说明。

1. tRNA 的二级结构

tRNA 分子只由一条链组成，含有 73~94 个核苷酸，其中有不少是修饰的核苷酸或修饰的碱基（如次黄嘌呤、硫尿嘧啶和假尿苷等），链内的大多数碱基通过氢键相连，但几乎所有的 tRNA 分子上不变

的核苷酸都位于三叶草结构上的非氢键区域。构成 tRNA 二级结构的要素有：环(loop)、茎(stem)和臂(arm)。一个典型 tRNA 的二级结构像三叶草(cloverleaf)，含有 4 个环和 4 个茎。环是由链内没有配对的碱基突出而成，茎是链内互补的碱基之间配对形成的局部 A 型双螺旋，臂则是紧靠着茎又不属于环的非配对核苷酸。

按照从 5′→3′ 的顺序，4 个环依次是 D 环(D loop)、反密码子环(anticodon loop)、可变环(the variable loop)和 TψC 环(图 6-23)。4 个茎依次是受体茎(acceptor stem)、D 茎、反密码子茎和 TψC 茎。臂有 D 臂、反密码子臂、TψC 臂和氨基酸臂。

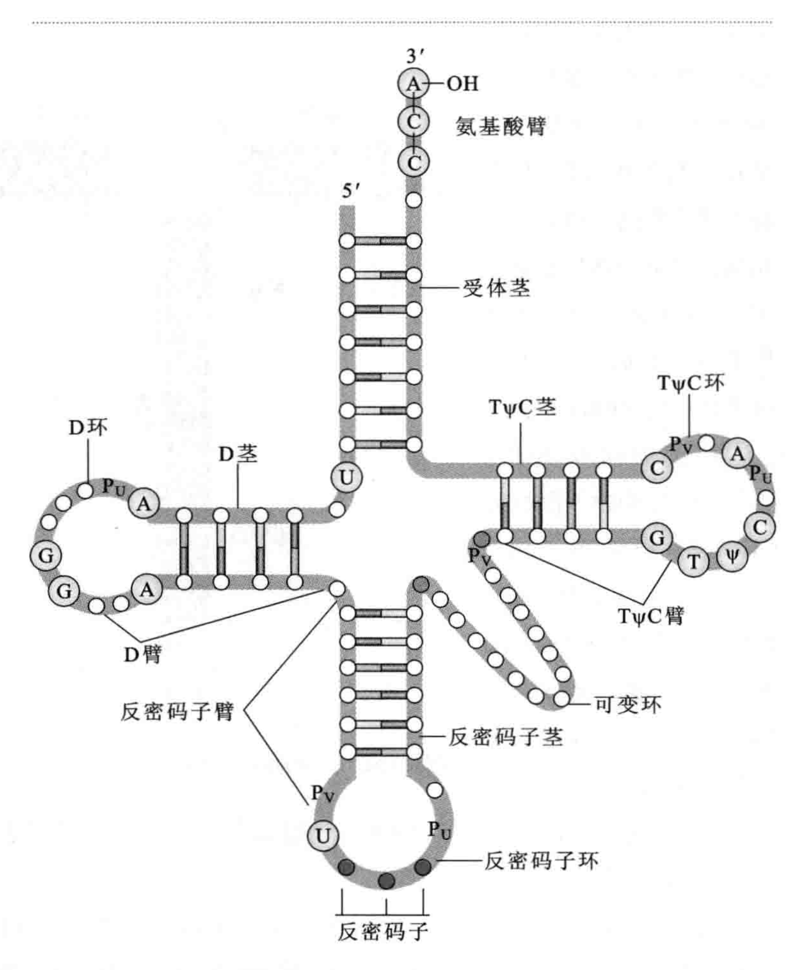


图 6-23 tRNA 的二级结构

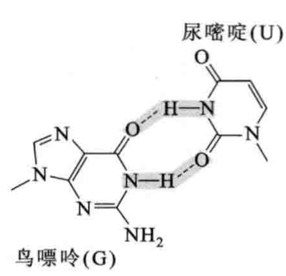


图 6-22 RNA 分子中的 GU 碱基对

受体茎是紧靠氨基酸与 tRNA 连接形成氨酰-tRNA 的地方，由 tRNA 靠近两端的互补序列配对而成。虽然不同的 tRNA 在一级结构上差别很大，但它们的二级结构却惊人地相似，并且 3′端的最后三个核苷酸总是 CCA，与 3′端的第四个核苷酸一起，并不参与形成受体茎，而是构成接受氨基酸的臂。细胞在合成氨酰-tRNA 的时候，氨基酸最终被添加到 CCA 末端腺苷酸的 3′-羟基上，在后续蛋白质生物合成中作为氨基酸残基的供体。D 环得名于它常常含有一个修饰的碱基，即二氢尿嘧啶(dihydrouridine，D)。反密码子环因含有反密码子而得名。反密码子是由 3 个核苷酸组成的单位，它能够按照碱基互补配对的原则去阅读 mRNA 上同样由 3 个核苷酸组成的特定密码子，从而将核苷酸序列翻译成氨基酸序列。紧靠反密码子 5′端的核苷酸总是 U，紧靠 3′端的核苷酸总是嘌呤核苷酸(通常烷基化)。在反密码子茎和 TψC 茎之间通常还有一个可变环或附加环(extra loop)，其长度在不同的 tRNA 分子上会有变化，因此有时可以用它来区分不同的 tRNA。TψC 环含有 7 个没有配对的碱基，包括 TψC 序列，其中 ψ 即假尿苷。核糖体与 tRNA 的结合依赖于其对 TψC 环的识别。

Quiz9 如果一个 tRNA 分子上的反密码子序列是 GψC，那么它识别的密码子序列是什么？

2. rRNA 的二级结构

核糖体是蛋白质生物合成的场所，构成核糖体的 RNA 被称为核糖体 RNA(ribosomal RNA，rRNA)。根据沉降系数大小，rRNA 可分为几种不同的类型。原核生物有 5S rRNA、16S rRNA 和 23S rRNA；真核生物有 5S rRNA、5.8S rRNA、18S rRNA 和 28S rRNA。在所有的 rRNA 分子上都发现有大量链内互补的序列，这些序列通过互补配对，使 rRNA 高度折叠。在不同物种的同一类型的 rRNA 上存在十分保守的折叠样式。

以 16S rRNA 为例，其内部存在大量的短螺旋，螺旋之间夹杂着多种形式的环和突起，这是根据形成氢键的序列比对(alignment)出来的。

比较不同物种来源的 16S rRNA 的一级结构和二级结构发现，尽管它们在一级结构上相似度并

不高，但它们的二级结构却惊人地相似。显然，16S rRNA 的分子进化是二级结构在起作用，而不是表现在核苷酸的序列上。换句话说，16S rRNA 在进化中，只要保证它的二级结构基本不变就行，而不用在乎其一级结构的变化。

再以 5S rRNA 为例，它的二级结构由 5 段螺旋（Ⅰ～Ⅴ）、1 个突环（环 A）、2 个发夹环（环 C 和环 D）、2 个内部环（环 B 和环 E）和 1 个铰链区组成（图 6-24）。图中的内部环 E 并没有画成环，是因为该环的 5′端序列和 3′端序列在三级结构上通过非 Watson-Crick 碱基对结合在一起。

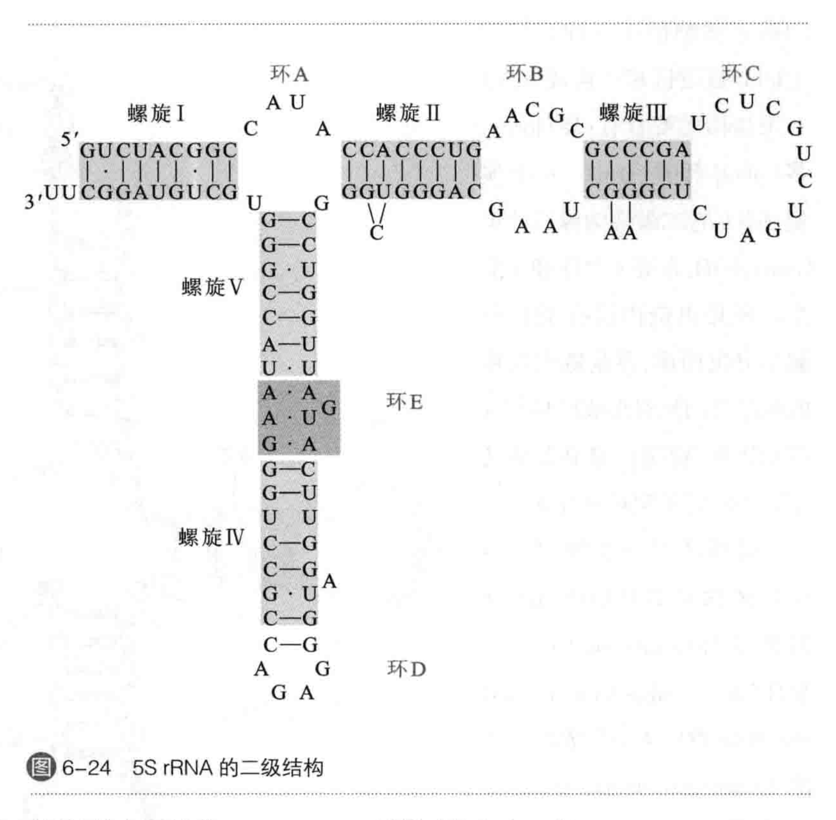


图 6-24　5S rRNA 的二级结构

3. mRNA 的二级结构

mRNA 的种类繁多，对各种 mRNA 的二级结构的研究却很少。实际上，人们关心更多的是它们的一级结构，这是因为决定多肽或蛋白质氨基酸序列的是其一级结构。然而，研究发现，mRNA 分子的二级结构，尤其是两端的二级结构，对翻译有一定的影响，某些 mRNA 正是借助于末端特殊的二级结构，对基因的表达进行调控。出现在 mRNA 分子上最多的二级结构部件也是茎环结构。

4. DNA-RNA 杂交双链的二级结构

由 DNA 和 RNA 形成的杂交双链形成的是 A 型双螺旋。这样的结构会在体内 DNA 复制、转录和逆转录的时候出现。在 DNA 复制的时候，先要合成一段短的 RNA 引物，这段 RNA 序列与 DNA 一条链上的部分序列是互补的，它们之间形成局部的 A 型双螺旋（参看第三十三章“DNA 复制”）；在基因转录的时候，以 DNA 一条链上的一段核苷酸序列作为模板合成各种 RNA 分子，被转录的 RNA 分子总是有 7~8 个核苷酸与 DNA 的模板链以 A 型双螺旋结合；而在逆转录的时候，RNA 是模板，合成的 cDNA 与其也是形成杂交双螺旋结构。

（二）RNA 的三级结构

细胞里的绝大多数 RNA，特别是较为稳定的 RNA，如 rRNA 和 tRNA 等，都有自己独特的三级结构，它们的三级结构也是通过折叠实现的。RNA 的三级结构是在二级结构的基础上进一步折叠、包装而成的，其复杂性不亚于蛋白质。其中的双螺旋区域主要充当刚性的框架结构来组织其他结构或功能部件。构成突起、内部环、发夹环和末端环的单链区域对于 RNA 最终三级结构的形成至关重要，其作用相当于氨基酸残基的侧链基团对于蛋白质三级结构形成的贡献。正是因为这些单链区之间以及单链区与双链区之间核苷酸的相互作用，才使得 RNA 最终能够折叠成它所特有的三级结构。事实上，RNA 折叠，特别是较大的 RNA，就像大多数蛋白质的折叠一样，也需要分子伴侣的帮助，以便让大多数 RNA 能快速折叠成正确的构象，而不至于陷在许多可能错误的构象中“不能自拔”。细胞内有许多非特异性的核酸结合蛋白（nonspecific nucleic acid-binding protein，NABP）充当 RNA 折叠的分子伴侣，它们通过防止或解除错误的折叠来指导一种 RNA 正确的折叠。为了与参与蛋白质折叠的分子伴侣区分开来，通常将帮助 RNA 折叠的分子伴侣称为 RNA 伴侣（RNA chaperone）。RNA 伴侣与正在折叠

的 RNA 分子结合以后，一旦出现错误的折叠，它就会破坏错误折叠的 RNA 分子内部的次级键，形成 RNA 蛋白体复合物，在这种复合物里，RNA 能迅速重新折叠成最终正确的构象。

以图 6-25 所示的枯草杆菌体内的 TPP 核开关为例，它在折叠的时候，先形成局部的二级结构，这些局部的二级结构的形成是相对独立的，主要利用氢键和碱基堆积力来驱动和稳定折叠。然后，在不同的二级结构之间发生远距离的相互作用，在此基础上形成构建三级结构的模块即模体。参与远程作用的主要是游离在双螺旋结构之外的环和突起中的核苷酸，这些核苷酸之间发生相互作用，形成非标准的碱基对或发生碱基堆积，在多数情况下，核糖 2′-OH 参与形成氢键。有时，相邻的螺旋之间可通过共轴的碱基堆积（coaxial base stacking），形成一段连续的或准连续的共轴螺旋（coaxial helix）。

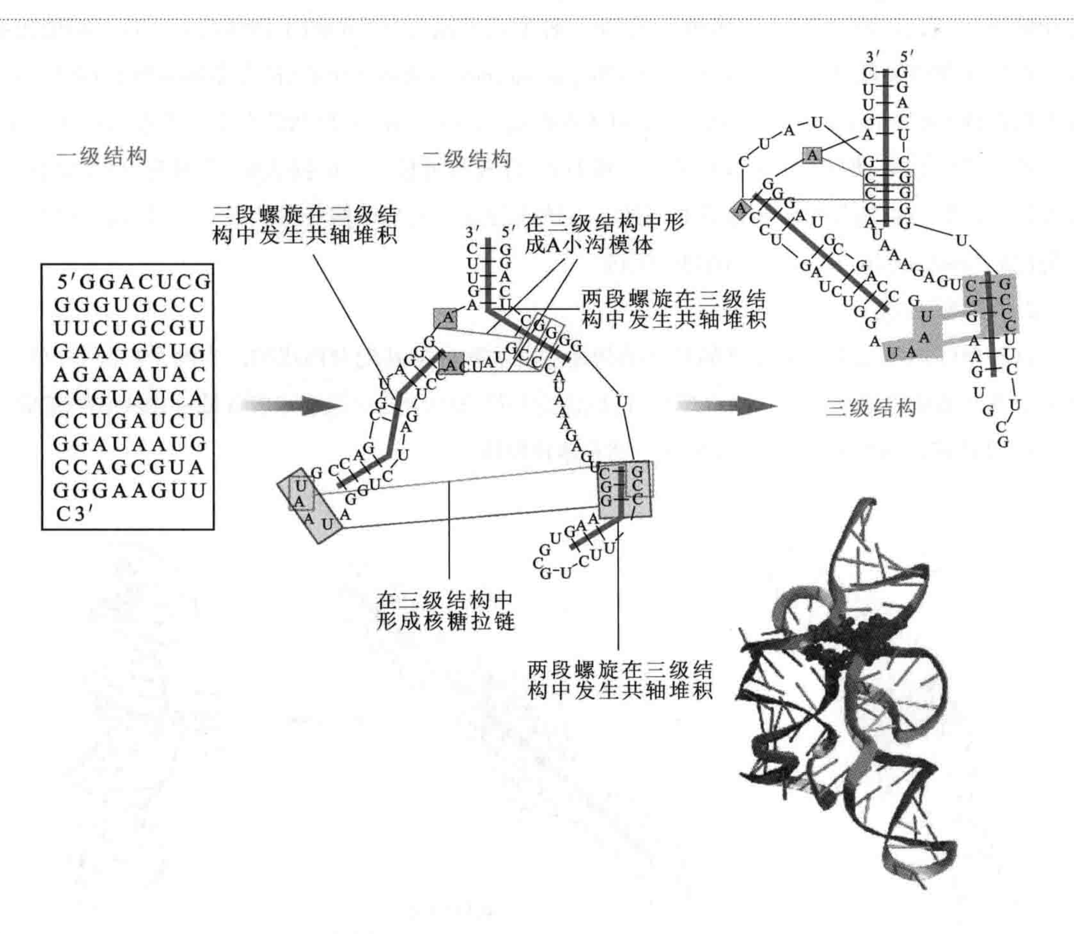


图 6-25 枯草杆菌 TPP 核开关的三级结构的形成

驱动和稳定 RNA 三级结构形成的因素经常涉及金属离子（如 Mg^{2+}）和碱性蛋白，这是因为 RNA 链在生理 pH 下本身是带高度负电荷的，需要通过与金属离子或者碱性的蛋白质的结合，来中和或屏蔽主链上磷酸基团带有的负电荷，以使不同区域的磷酸核糖骨架能相互靠近，发生近距离的接触和包装，并可能作为最终构象的一部分。例如，I 组内含子核酶就围绕 1 个 Mg^{2+} 进行折叠。此外，其他几个因素也能影响到 RNA 的折叠，包括：在一级结构上相距较远的两段区域形成标准的 Watson-Crick 碱基对和非 Watson-Crick 碱基对，在一段双螺旋和一段单链之间通过非 Watson-Crick 碱基对形成三螺旋（图 6-26），以及碱基和碱基之间的堆积力、碱基和主链之间（特别是核糖部分）的相互作用等。

Quiz10 为什么 RNA 的三维结构也比 DNA 复杂得多?

构成 RNA 三级结构的主要结构模体有多种类型，如假节结构（pseudoknot）、“吻式”发夹结构（kissing hairpin）。这些不同的模体之间可以相互作用，形成更复杂的结构。

1. 假节结构

这是 RNA 分子上最常见的一种模体，最初是在萝卜黄色镶嵌病毒（turnip yellow mosaic virus）的基因组 RNA 上被发现的。一个假节结构至少是由两段螺旋和将两段螺旋联系起来的单链区域或环组成（图 6-27）。目前在几种拓扑学结构不同的假节结构中，性质最为确定的是 H 形假节。在 H 形假节

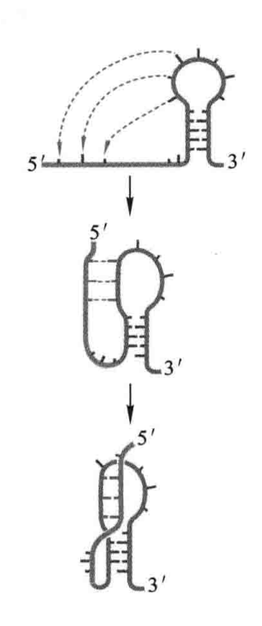


图6-27 假节结构的形成(Watson 等,2007)

结构之中,一个发夹环上的碱基与茎以外的碱基形成分子内的配对,从而形成第二个茎环结构,产生具有两茎、两环的假节结构。上述两茎能够相互垛叠在一起,形成一个具有一个连续股和一个不连续股的准共轴螺旋。单链环区域经常与相邻的茎发生作用,形成氢键,参与整个分子结构的形成。因此,这种相对简单的折叠能产生既复杂又稳定的结构。由于环和茎长度以及它们之间相互作用的变化,假节结构实际上有多种形式,每一种形式可能具有不同的生物学功能。这些功能包括:参与形成多种核酶和自我剪接的内含子(self-splicing intron)的活性中心;诱导多种病毒在翻译过程中发生核糖体移框(ribosomal frameshifting);对于端粒酶(telomerase)发挥活性有十分重要的作用。虽然端粒酶不属于核酶,但其 RNA 部分充当端粒 DNA 合成的模板。有证据表明,端粒酶 RNA 中有一个高度保守的假节结构是端粒酶活性必需的。人体内的这个假节结构若发生突变,可导致一种叫先天性角化症(dyskeratosis congenita)的遗传疾病。

U:A:U三联体

图6-26 在三螺旋结构中的三联体配对

2. "吻式"发夹

这种模体结构是由两个独立的发夹结构通过环之间的碱基配对形成的。当两个环配对以后,就在两个发夹结构之间形成第三段双螺旋,这段螺旋与原来的两端双螺旋形成共轴堆积或共轴螺旋(图6-28)。艾滋病病毒的基因组 RNA 分子上就有这种模体。

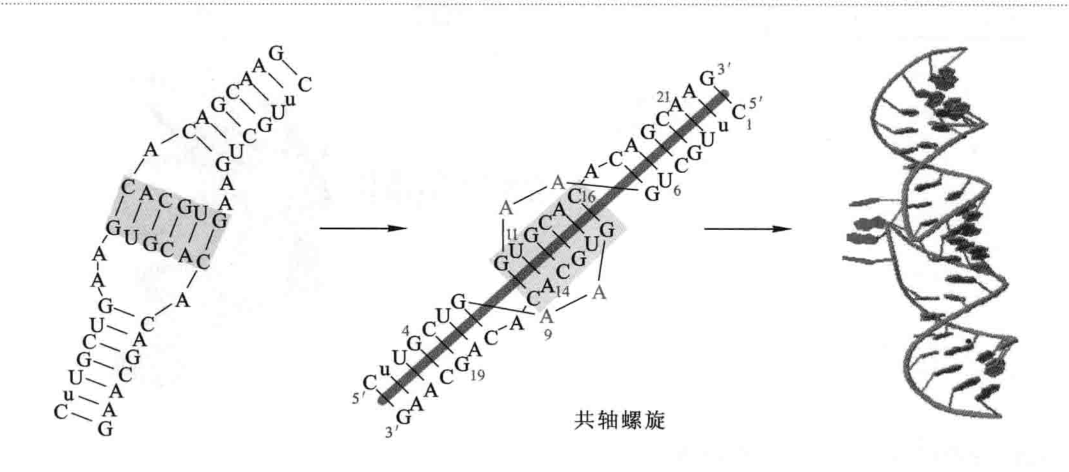


图6-28 艾滋病病毒基因组 RNA 分子中吻式发夹结构的形成

除了假节结构和"吻式"发夹结构以外,现在还发现了 A- 小沟模体(A-minor motif)、核糖拉链(ribose zipper)和弯曲 - 转角(kink-turn)等。这些模体出现在不同类型的 RNA 分子上,赋予了对应的 RNA 特殊的结构和功能。

e6-3 RNA 分子上的其他结构模体

就 tRNA 分子而言,其三级结构的形成依赖于 D 环上的碱基和不变碱基以及 TψC 环上的碱基之间建立氢键。参与三级结构形成的许多氢键并不是通常的 AU 和 GC 碱基对。所有的氢键将 D 臂和 TψC 臂折叠到一起,并将三叶草二级结构弯曲成稳定的倒 L 型(图 6-29)。

在三级结构中,tRNA 的两个功能端被有效地隔离开,携带氨基酸的受体茎位于 L 的一端,与另一端的反密码子相距 7 nm 左右,而 D 环和 TψC 环构成 L 的角。在倒 L 形构象中,碱基的排列方向都尽可能增加碱基平面之间的疏水堆积力,这也是仅次于氢键的稳定倒 L 形构象的因素。

对 rRNA 的三级结构而言,只是在近十多年来才有了许多突破。X 射线衍射获得的数据表明,核糖体的整体构象是由 rRNA 决定,核糖体蛋白质一般正好位于 RNA 螺旋之间,起点缀作用,图 6-30 分别显示了原核生物核糖体小亚基和大亚基上 rRNA 的三维结构。

图 6-29 tRNA 的倒 L 形三级结构以及三级氢键的配对（Garrett 等，2010）

第四节 核酸与蛋白质的相互作用

细胞内的核酸并不是游离的，它们总是与蛋白质或酶发生各种各样的作用，这些作用不但能够影响核酸的结构，而且直接参与基因的复制、重组、修复、转录、转录后加工和翻译等过程。在很多情况下，核酸与蛋白质能够形成紧密的复合物即核酸蛋白体颗粒。下面就几种主要核酸蛋白体的结构与功能进行简单的介绍，至于在 DNA 复制、转录、重组和修复中 DNA 与蛋白质的相互作用，将在第三篇“分子生物学”中做详细的阐述。



小亚基 rRNA



大亚基 rRNA

图 6-30 原核生物核糖体大、小两个亚基 rRNA 的三维结构

一、DNA 与蛋白质的复合体

DNA 与蛋白质形成的复合体最为重要的是染色质（chromatin）或染色体（chromosome）以及 DNA 病毒颗粒，这里仅介绍染色质和染色体。

（一）染色质

与细菌基因组形成鲜明对照的是，真核生物的基因组 DNA 是高度折叠的，并与一些特殊的蛋白质结合形成一种致密的动态多聚物，即染色质。这种蛋白质 –DNA 复合物在细胞核内发生的 DNA 复

制、重组、修复和转录过程中起重要的作用。

染色质是一种动态的结构，通常在细胞核内被高度压缩成无定形的长纤维。在有丝分裂的中期呈高度浓缩。染色质的组成包括 DNA、蛋白质和少量的 RNA。在质量上，DNA 和紧密结合的蛋白质差不多各占一半。染色质上的蛋白质有组蛋白（histone）和非组蛋白两类（non-histone protein）。

染色质可分为常染色质（euchromatin）和异染色质（heterochromatin）。常染色质在细胞周期的间期浓缩程度很低，染色淡，具转录活性；异染色质浓缩程度高，染色深，一般无转录活性。异染色质 DNA 在 S 期复制时比常染色质要晚，这是异染色质高度浓缩的结果。

染色质中最重要的结构单元是核小体，核小体是由 DNA 和组蛋白组成的串珠状（beads on a string）结构。

1. 组蛋白

组蛋白是染色质上含量最丰富的蛋白质，富含碱性氨基酸——Lys 和 Arg，属于一类较小的碱性蛋白。它们在生理 pH 下带有正电荷，这样就可以和带负电荷的 DNA 通过静电引力结合在一起。

组蛋白主要有 5 种类型，即 H1、H2A、H2B、H3 和 H4。某些生物或组织还有 H1° 或 H5（表 6–4）。此外，细胞内还存在与这 5 种标准形式相对应的不同变体。这些变体在取代相应的标准组蛋白与 DNA 结合以后，会改变局部的核小体和染色质的结构，进而有可能影响到基因的表达（参看第四十三章“真核生物的基因表达调控”第一节“在染色质水平上的调控”）。

表 6–4 几种组蛋白的性质比较

组蛋白	M_r	保守性
H3	15 400	高度保守
H4	11 340	高度保守
H2A	14 000	在不同组织和物种中，中度保守
H2B	13 770	在不同组织和物种中，中度保守
H1	21 500	在不同组织和物种中，显著变化
H1°	~21 500	变化很大，只存在于非复制的细胞
H5	21 500	高度变化，只存在于某些物种无转录活性的细胞

各种组蛋白在进化的保守性上是不一样的，保守性最高的是 H4，其次是 H3，再其次是 H2A 和 H2B。变化最大的是 H1，但某些组织中没有 H1，而含有其他类型的组蛋白。例如，在鸟类的红细胞中由 H5 取代了 H1。

在三维结构上，H2A、H2B、H3 和 H4 的结构相似，N 端形成尾巴，C 端形成组蛋白特有的结构模体——组蛋白折叠（histone fold）。组蛋白折叠由 3 段 α 螺旋组成，中间的 α 螺旋比较长，两侧的较短，组合起来形如一个浅的“U”字（图 6–31）。

组蛋白氨基酸序列的保守性清楚地表明，它们在各种真核生物的染色质中所起的作用十分相似。

2. 非组蛋白

非组蛋白泛指在染色质上除组蛋白以外的各种蛋白质。其种类繁多，功能广泛，成分也不固定，许多尚未定性。

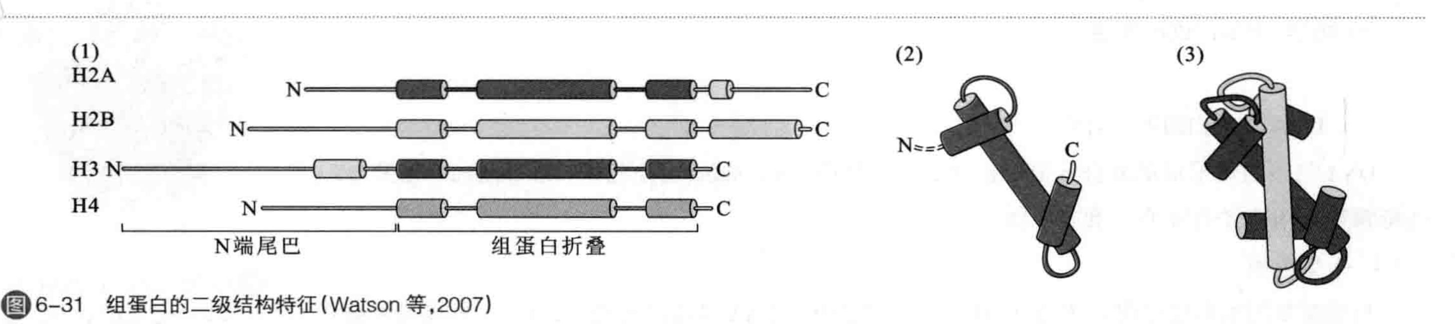


图 6–31 组蛋白的二级结构特征（Watson 等，2007）

3. 核小体

核小体(nucleosome)可视为染色质的一级结构单位,在电镜下呈串珠状(图 6–32):每一个"珠子"由组蛋白核心(histone core)和环绕其上(而不是穿过)的 DNA 组成(约 146 bp);相邻"珠子"之间的连线为 DNA,称为连接 DNA(linker DNA),长度 8~114 bp 不等,它最容易受到 DNA 酶的水解。H1 与连接 DNA 结合,但去除 H1 并不会破坏核小体结构(图 6–33)。

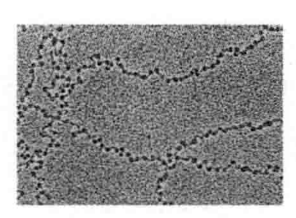
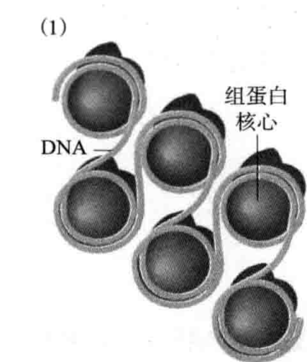


图 6–32 电镜下的核小体结构

组蛋白核心是一个八聚体,由 4 组二聚体通过组蛋白折叠结合在一起。H3 与 H4 通过一个折叠形成异源二聚体,H2A 与 H2B 通过另一个组蛋白折叠形成另一个异源二聚体。两个 H3–H4 二聚体通过 H3 之间的四螺旋束(4–helix bundle)形成 H3–H4 四聚体,最后一对 H2A–H2B 通过 H2B 与 H4 之间的相互作用形成八聚体。组蛋白核心通过静电引力与 DNA 结合,其表面大约环绕 146 bp 的 DNA 双螺旋,因此 DNA 长度被压缩 6~7 倍。组蛋白也与一些参与组蛋白核心组装和调节基因表达的非组蛋白结合。有时组蛋白还会受到乙酰化、甲基化和磷酸化等形式的共价修饰,以改变分子的电荷,从而影响到它们与 DNA 的相互作用。

最新研究发现,细胞内核小体的组装经历前核小体(pre-nucleosomes)中间物。这种介于 DNA 与核小体之间的前核小体结构,在依赖于 ATP 的马达蛋白(ATP–dependent motor protein)ACF 的催化下,可快速地形成标准的核小体结构。

(1)



(2)



(3)

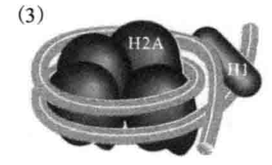


图 6–33 核小体结构模型

核小体的宽度约为 10 nm,因此这个阶段的染色质被称为 10 nm 纤维或核蛋白纤维(nucleoprotein fibril)。通过 X 射线晶体衍射分析发现,构成核小体核心的组蛋白单体 N 端和 C 端尾巴并不在核心结构之中,而是伸出来通过超螺旋上的沟与相邻的核小体接触;DNA 双螺旋每隔 10 bp 其小沟就面对蛋白质的表面,并与 Arg 侧链接触;DNA 的大沟朝外,能够被序列特异性调节蛋白识别并结合;核小体表面的 DNA 卷曲并略显缠绕不足,以负超螺旋的形式存在(螺距为 10.2 bp/ 圈而不是 10.5 bp/ 圈)。

(二) 染色体

染色体是基因组的结构单位。每一条染色体由一个 DNA 分子和与之相结合的蛋白质组成,一般呈高度浓缩的状态。

细菌含有一个单一的连续环状染色体,呈超螺旋状态,被包装成多个来回的环,没有核膜包围,但集中在某一区域,并与质膜接触。出现在细菌染色体上的蛋白质主要是一些碱性蛋白。其中最普遍的是 HU 蛋白。细菌没有组蛋白,因此不会形成核小体的结构。

与细菌相似,古菌的染色体也是单一的连续环状染色体,但许多古菌具有组蛋白,因此也会形成核小体。然而,古菌的组蛋白要短于真核生物,而且在与 DNA 形成核小体的时候,只形成四聚体核心。如此小的组蛋白核心让古菌的一个核小体只能包被约 80 bp 的 DNA。

真核生物有多个线形染色体。不同的染色体大小不一样,但都呈高度浓缩的状态。每一条染色体都含有 3 个重要的功能元件:自主复制序列(autonomously replicating sequence,ARS)、着丝粒(centromere)和端粒(telomere)。这些元件是线形染色体正常复制及其随后的正确分离所必需的。其中 ARS 充当 DNA 复制起始区,着丝粒的主要功能是使复制的染色体在细胞有丝分裂和减数分裂过程中可均等地分配到子细胞中,端粒为线性 DNA 的两端,是维持染色体的完整性所必需的。

染色体是在 DNA 与组蛋白形成核小体的基础上,经过层层包装形成的,先后经历了 30 nm 纤维 / 螺线管、环、"玫瑰花瓣"等结构,直至形成高度浓缩的染色体。

e6–4 真核生物染色体的包装

二、RNA 与蛋白质的复合体

细胞内的 RNA– 蛋白质复合体很多,有 RNA 病毒、核糖体、信号识别颗粒、剪接体、编辑体、核糖核酸酶 P 和端粒酶等。它们的结构与功能将会在以后的相关章节中一一介绍。

第五节　核酸的功能

早在1868年，孟德尔遗传规律被发现之后不久，瑞士的J. T. Miescher就发现细胞核含有一种富含磷的酸性物质，定名为核素(nuclein)，后来被称为核酸。但此项重大发现与遗传规律的发现一样遭受了相同的命运，当时没有引起人们的重视，直到20世纪30年代末，人们才逐渐将核酸和细胞的功能联系起来。

对于RNA而言，其功能多种多样。在远古的"RNA世界"中，RNA可能行使过当今DNA和蛋白质承担的所有功能，而在现代的生命世界中，RNA的生物功能仍然还有很多，主要包括：①充当RNA病毒的遗传物质，如甲肝病毒和艾滋病病毒；②作为生物催化剂即核酶，如核糖核酸酶P和核糖体；③参与蛋白质的生物合成，这与细胞内3种最重要的RNA，即mRNA、tRNA和rRNA有关；④作为引物，参与DNA复制(参看第三十三章"DNA复制")；⑤参与RNA前体的后加工，如snRNA参与细胞核mRNA前体的剪接，snoRNA参与真核rRNA前体的后加工，gRNA参与编辑(参看第三十七章"转录后加工")；⑥参与基因表达的调控，如干扰RNA和反义RNA参与翻译水平上的基因表达调控；⑦参与蛋白质共翻译的定向和分拣，如7S RNA(参看第四十章"蛋白质的翻译后加工及其定向和分拣")；⑧参与X染色体的失活，这与Xist RNA有关。

与RNA相比，DNA在生物体内的功能只有一种，就是作为生物体的主要遗传物质，生物体的一切性状最终都是由它决定的。证明这项功能的主要证据是O. T. Avery的细菌转化实验和Hershey-Chase的噬菌体实验。

e6-5 细菌转化实验和噬菌体实验

科学故事——DNA双螺旋结构的发现

离DNA双螺旋60岁的生日越来越近了！不过双螺旋结构的共同发现者之一Crick已于2004年7月29日病逝，享年88岁。

1953年2月28日中午，剑桥大学两位年轻的科学家Crick和Watson在一个非学术场所，位于剑桥大学国王学院斜对面的老鹰酒吧(The Eagle Pub)宣布他们的发现：DNA是由两条互补的脱氧核苷酸链组成的双螺旋结构。1953年4月25日，*Nature*发表了这一成果。不过，按照国际学术界的惯例，一项成果必须是在学术杂志上正式发表才能被视为正式宣布，这样做是为了防止有人钻空子随便宣布获得重大成果而造成混乱。因此，尽管Watson和Crick于2月28日就在老鹰酒吧宣布了这一成果，但包括英国官方机构在内的很多机构仍然把4月25日作为DNA双螺旋结构的发现纪念日。

与Watson和Crick分享诺贝尔奖的还有Maurice Wilkins。Wilkins的贡献在于为Watson和Crick的发现提供了实验证据。然而科学界就究竟是谁第一个发现了DNA的双螺旋结构一直存在不同的看法。眼下，一些DNA双螺旋结构发现者的私人手稿和信件揭示了科学史上一些鲜为人知的故事与细节。这可以算是科学史上的一桩著名公案，其中涉及另外一位著名的科学家Rosalind Franklin。Franklin是一位非常优秀的实验科学家。在与Wilkins合作期间凭着独特的思维，设计了能从多方面了解DNA结构的实验方法，获取在不同温度下的DNA纤维的X射线衍射图，从而分辨出了这种分子的维度、角度和形状。她发现DNA是螺旋结构，至少有两股，其化学信息面朝里。这已经非常接近真理。然而，Franklin非常有个性，经常对人进行直言不讳的尖锐批评，与Wilkins因性格不合也时常发生矛盾，最终她从Wilkins小组中分离了出来，另立门户。

受Wilkins和Franklin关于DNA的X射线衍射图分析报告的启发，在英国的卡文迪什实验室工作时Watson与Crick相遇并共同研究DNA的结构。最初，Watson与Crick千辛万苦地按照他们的理解把糖和磷酸构成的主链置于中间，碱基则朝外，搭出了DNA三螺旋的结构。他们认为，这个模型与Wilkins和Franklin提供的X射线衍射图比较吻合。在向Wilkins和Franklin透露了他们的所谓最新重大成果时，Franklin一针见血地指出了这一成果的缺陷，特别是这个模型过分模仿水分子。尽管Franklin当时并不知

道 DNA 的精确结构应当是什么样的，但是通过她自己的研究，她至少知道 DNA 结构不应当是什么样的。也就是说，DNA 的螺旋结构并不是三螺旋。正是她这种独特的指路明灯式的光芒，把 Watson 和 Crick 一步步引导到了正确的方向。

作为一名女犹太裔科学家，再加上脾气率直，Franklin 自然不被学术界所包容。因此 1962 年，Watson 和 Crick 获得诺贝尔奖时发表演说根本没有提到她。

然而 Franklin 的贡献是毋庸置疑的，她分辨出了 DNA 的两种构象，并成功地拍摄了它的 X 射线衍射照片，特别是有一张编号 51 的照片格外清晰。Watson 在 1968 年出版的 *The Double Helix* 一书中，透露了 Wilkins 曾偷偷复制 Franklin 的研究成果并提供给他，其中就包括了现在众所周知的她证明螺旋结构的第 51 张 X 射线图像。Watson 和 Crick 未经她的许可使用了这张照片。在一封 Franklin 与 Watson、Crick 的通信中，Crick、Watson 对 Franklin 说，她和 Wilkins 的非常清晰的 DNA 图片对他们启发很大。可以说，如果没有 Franklin 的 X 射线衍射成果，要确定 DNA 的螺旋结构是很难的。

这个故事的结局是伤感的。当 1962 年 Watson、Crick 和 Wilkins 获得诺贝尔奖的时候，Franklin 却于 1958 年因卵巢癌去世了，年仅 37 岁。而按诺贝尔立下的规矩：诺贝尔奖只发给那些尚在人世、并为人类和社会发展做出了极大贡献的人，因而 Franklin 未能获得该殊荣。

e6-6 DNA 双螺旋的发现（全版）

e6-7 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/nucleic-acids.html（完全免费的医学生物化学课程网站，有关核苷酸和核酸的内容）
2. http://en.wikipedia.org/wiki/DNA（维基百科有关 DNA 的内容）
3. http://en.wikipedia.org/wiki/RNA（维基百科有关 RNA 的内容）
4. http://www.blc.arizona.edu/molecular_graphics/dna_structure/dna_tutorial.html（美国亚利桑那大学另一生物化学课程网站，内有 DNA 基本结构的介绍）
5. http://www.tulane.edu/~biochem/nolan/lectures/rna/intro.htm（美国杜兰大学的一个生物化学课程网站，内有核酸结构学习辅助资料，很实用）

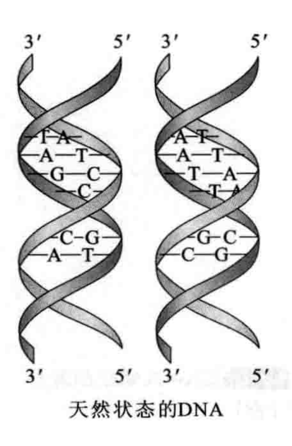
参考文献：

1. Zhao Keji, et al. SWI/SNF-mediated chromatin remodeling induces Z-DNA formation on a nucleosome. *Cell & Bioscience*, 2(1):3, 2012.
2. Kadonage J T, et al. Identification of a rapidly formed nonnucleosomal histone-DNA intermediate that is converted into chromatin by ACF. *Molecular Cell*, 43(4):638-648, 2011.
3. Nielsen P E, Egholm M. An introduction to peptide nucleic acid. *Current Issues Molec Biol*, 1(2):89-104, 1999.
4. Wells R A, et al. The chemistry and biology of unusual DNA structures adopted by oligopurine-oligopyrimidine sequences. *FASEB J*, 2:2 939-2 949, 1988.
5. Rich A, et al. Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, 282:680-686, 1979.
6. Watson J D, Crick F H C. Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature*, 171:737-738, 1953.

第七章　核酸的性质及研究方法

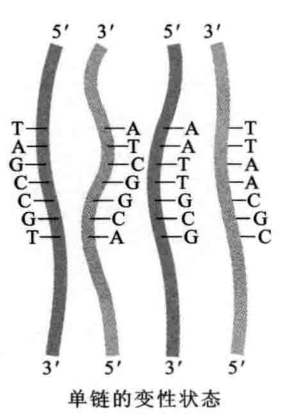
核酸是由核苷酸组成的多聚物，因此所具有的许多性质是由核苷酸残基赋予的，例如紫外吸收和两性解离。但作为生物大分子的核酸又有许多特有的性质，例如变性和复性。

本章将主要介绍核酸的理化性质，以及如何利用这些性质对核酸进行分离、纯化，还会对核酸的一些研究方法做简单的介绍。



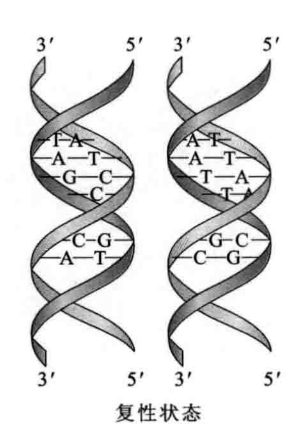
天然状态的DNA

加热或碱　变性



单链的变性状态

特殊的条件下　复性



复性状态

图 7–1　DNA 的变性和复性

第一节　核酸的理化性质

核酸的理化性质主要包括紫外吸收（参见第五章“核苷酸”）、酸碱解离、变性、复性和水解等。

一、酸碱解离

核酸的酸碱解离性质与核苷酸一致，详见第五章“核苷酸”。但由于核酸为多聚核苷酸，含有大量的磷酸基团，因此它们的 p*I* 较低，DNA 的 p*I* 为 4~4.5，RNA 的 p*I* 为 2~2.5。

二、变性

核酸的变性是指核酸受到加热、极端的 pH、离子强度的降低或特殊的化学试剂（尿素、盐酸胍和甲醛等）的作用后，其双螺旋区因氢键和碱基堆积力被破坏而发生解链、变成单链的过程（图 7–1）。其中并不涉及共价键的断裂。

核酸在变性时，其理化性质会发生一系列改变，例如紫外吸收和浮力密度升高，黏度降低，沉降速率加快，生物活性降低、丧失或不变，其中紫外吸收增加的现象称为增色效应（hyperchromic effect）。增色效应产生的原因是双螺旋结构之中的碱基堆积力降低了紫外吸收，当核酸变性以后，碱基堆积力被削弱，这时每一个碱基的紫外吸收能力都充分表现出来，紫外吸收随之升高。上述性质的变化可作为检测核酸变性的指标。

在此有必要将蛋白质变性和核酸变性进行比较。两者都强调了高级结构的破坏和次级键的断裂，且共价键并不受影响。然而，蛋白质变性强调的是蛋白质三级结构的破坏和生物功能的丧失。如果蛋白质发生变性，其三级结构肯定遭到了破坏，它的生物学功能也必然丧失；而核酸的变性强调的是其二级结构即双螺旋结构的破坏，它与核酸的生物学功能是否丧失没有必然的联系。

对于 DNA 而言，变性并不会使其生物学功能丧失，反而有利于其生物学功能的发挥，这是因为 DNA 的生物学功能是贮存、复制及转录遗传信息，而遗传信息是贮存在一级结构之中的，DNA 在变性的时候，一级结构并没有遭到任何破坏。此外，无论是 DNA 复制还是转录，第一步都需要 DNA 发生解链，这实际上就是 DNA 的变性。例如，体外复制 DNA 即 PCR 的第一步反应就是热变性。

而对于 RNA 而言，变性是否破坏其生物学功能，需要区别对待。那些生物学功能直接由高级结构决定的 RNA，如 tRNA、rRNA、snRNA、snoRNA 和核酶，一旦发生变性，就会像蛋白质一样，生物学功能立刻丧失。而对于那些生物学功能直接与一级结构有关的 RNA，如 mRNA 和 RNA 病毒的基因组 RNA，变性不会破坏它们的生物学功能。

如图 7–2 所示，双链 DNA 的热变性是在很窄的温度内发生的，与晶体在熔点时突然熔化的情形

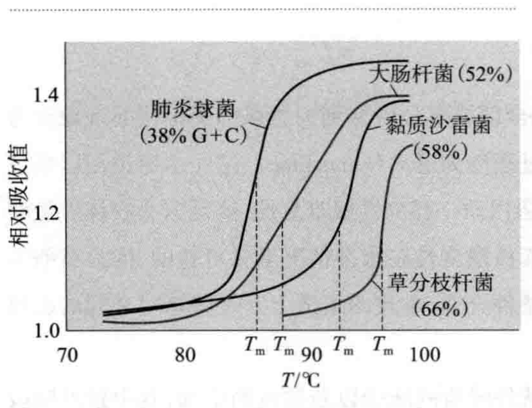


图 7-2　GC 含量对 DNA T_m 的影响

相似，因此 DNA 也具有“熔点”，用 T_m 表示。T_m 是指 DNA 双螺旋有一半发生热变性或有一半氢键因受热破坏时相应的温度。DNA 的 T_m 值通常在 82~95℃。

Quiz1 细胞内的 DNA 是如何发生变性的？

若以温度（T）对 DNA 溶液的紫外吸光度作图，得到的 DNA 变性曲线通常为 S 形。S 形曲线下方平坦段，表示 DNA 的氢键尚未破坏；当加热到某一温度，氢键突然断裂，DNA 迅速解链，同时伴随着吸光度的陡然上升，这对应于曲线中段陡直的部分；此后因“无链可解”而出现增色效应丧失的上方平坦段。如果需要强调热变性与增色效应之间的关系，那么可以从另外一个角度来定义 T_m，即让增色效应达到一半时的温度作为 T_m，它在 S 形曲线上相当于吸光度增加的中点处所对应的横坐标。

Quiz2 你如何解释双螺旋 DNA 变性的曲线为 S 型？

DNA 的 T_m 值并不是固定不变的，至少受到 4 种因素的影响：

(1) DNA 的均一性　DNA 的均一性有两种不同的含义：第一种是指 DNA 序列的均一性，如人工合成的 poly d(A–T) 或 poly d(G–C) 具有高度的均一性，这是因为它们只含有一种碱基对，与天然 DNA 相比，这些人工合成的高度均一性 DNA 的 T_m 值范围就很窄，这是因为它们在变性时的氢键断裂几乎同时进行，所要求的变性温度更趋于一致；第二种是指待测样品 DNA 的组成是否均一，即是否含有其他杂 DNA 的污染，例如，所测样品中只含有一种病毒 DNA，其 T_m 值的范围就相对较窄，若混有其他来源的 DNA，T_m 值范围就变宽。总的说来，某种 DNA 均一性越高，变性时 DNA 双螺旋之间的氢键断裂所需能量就接近，其 T_m 值范围就越窄。反之亦然。

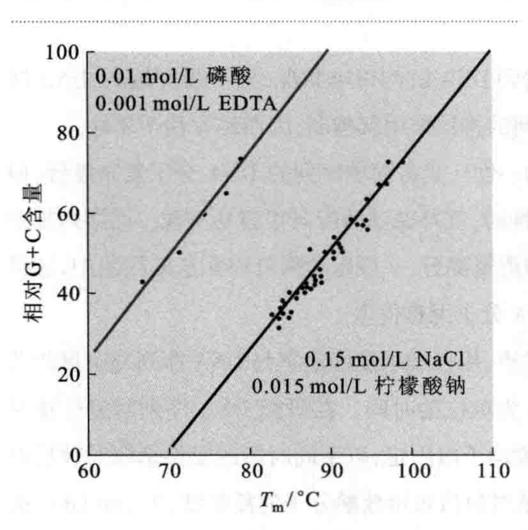


图 7-3　DNA 分子中的 GC 含量与 DNA T_m 之间的关系曲线

(2) GC 含量　在溶剂条件固定的前提下，T_m 值的高低取决于 DNA 分子中的 GC 含量。如图 7-3 所示，GC 含量越高，T_m 值越高。这是因为 GC 碱基对具有 3 个氢键，比 AT 碱基对多 1 个，而且产生的碱基堆积力更大。显然破坏 GC 碱基对比破坏 AT 碱基对需要消耗更多的能量，故 GC 含量高的 DNA，其 T_m 也高。实验证明，DNA 的 T_m 值与其 GC 含量呈正相关。

T_m 值与 GC 含量（w）百分比的这种关系可用以下经验公式来表示（DNA 溶解在 0.2 mol/L NaCl 溶液中）：$w\%(G+C)=2.44(T_m-69.3℃)$。

(3) 离子强度　溶液中的阳离子能够中和或屏蔽 DNA 主链上磷酸根的负电荷，减弱两条链之间的排斥作用而增强 DNA 双螺旋结构的稳定性。因此，溶液中的离子强度越高，DNA 的 T_m 值就越高。

(4) 其他变性因素的作用　许多容易形成氢键的试剂（例如尿素和甲酰胺）能够破坏碱基对之间的氢键，因此可以降低 DNA 的 T_m 值。而那些能够提高碱基溶解性的试剂（如甲醇），或者能够破坏 DNA 分子外被水化层的试剂（如三氯乙酸），因能减弱碱基之间的疏水作用，也会导致 T_m 的下降。某些蛋白质（如单链 DNA 结合蛋白）因为能够稳定 DNA 单链状态，也能降低 T_m 值。相反，许多蛋白质与 DNA 结合以后能够稳定双螺旋结构而阻止 DNA 变性，因此反而提高了 T_m 值。

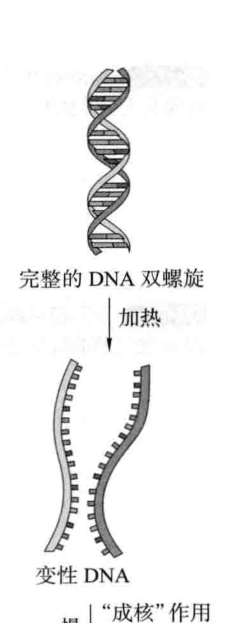
Quiz3 预测生活在南极与生活在火山口附近的微生物在基因组的一级结构和三级结构上会有什么不同？

RNA 的 T_m 值较为复杂，对于双链 RNA 来说，其 T_m 值的性质与 DNA 相近。但绝大多数 RNA 为单链，其分子内的双螺旋区域有限，导致 RNA 在变性时，性质变化程度不及 DNA，同时导致它的 T_m 值较低、变性曲线较宽。

三、复性和杂交

1. 复性

当各种变性因素消失时，变性时解开的互补单链全部或部分恢复到天然双螺旋结构的现象称为复性。热变性 DNA 一般经缓慢冷却后即可复性，此过程称为退火（annealing）。这一术语也用以描述杂交核酸分子的形成（见后）。与核酸变性不同的是，蛋白质一经变性很难复性，这是因为在体外重建与天然蛋白质一模一样的次级键几乎是不可能的。而核酸变性在很多情况下是可逆的，因为核酸变性以后，两条链之间的互补关系仍然存在，所以只要条件合适，重建两条链上互补碱基对之间的次级键是比较容易的（见图 7-1）。



伴随着 DNA 复性的是其浮力密度、沉降速率和紫外吸收的减少以及黏度的增加，其中紫外吸收减少的现象称为减色效应（hypochromic effect）。

DNA 复性的第一步是两个互补的单链分子间的接触以启动部分互补碱基的配对，这叫"成核"作用（nucleation）。随后，成核的碱基对经历小范围重排以后，单链的其他区域像"拉链"一样迅速复性（图 7-4）。

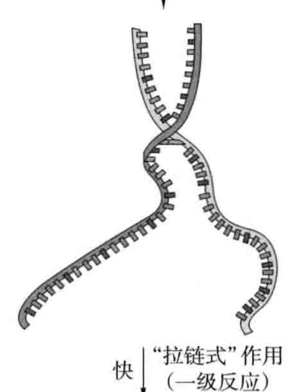




图 7-4 DNA 的复性历程（Garrett 等，2010）

影响 DNA 复性的因素有温度、DNA 浓度、离子强度和 DNA 序列的复杂度等。

（1）温度　一般认为低于 T_m 25℃左右的温度是复性的最佳温度，离此温度越远，复性速率就越慢。在很低的温度（如低于 4℃）下，分子的热运动显著减弱，互补链配对的机会自然大大减少。从分子热运动的角度考虑，维持在 T_m 以下较高的温度，实际上更有利于复性。此外，复性时温度的下降须缓慢进行，若在超过 T_m 的温度下迅速冷却至低温，复性几乎是不可能的。实验室中经常以此方式保持 DNA 的变性状态。

（2）DNA 浓度　DNA 浓度越高，则溶液中 DNA 分子越多，相互碰撞结合"成核"的机会越大，就越有利于复性。

（3）离子强度　DNA 溶液中的离子强度直接影响到 DNA 链的带电状况，离子强度越高，DNA 链上磷酸根基团被屏蔽的效果就越好，DNA 互补单链之间的排斥作用就越弱，因而越有利于复性。

（4）DNA 序列的复杂度（sequence complexity）或均一性　具有简单序列的 DNA 分子复杂度低，但均一性高，如 poly（dA）和 poly（dT）这两种单链序列复性时，互补碱基的配对很容易实现。而序列复杂的 DNA 复杂度高，但均一性低，如小牛胸腺 DNA 的非重复部分，一般以单拷贝存在于基因组中，这样的序列要完成互补配对，显然要比上述复杂度低的 DNA 分子困难得多。

Quiz4 在相同的条件下，序列 ATATATATATAT 和 ATAATAATAATA 与各自的互补序列发生复性，哪一种更容易？

在核酸复性动力学研究中，需要引入一个 *Cot* 的术语，用以表示复性速率与 DNA 序列复杂度的关系。其中 *Co* 为单链 DNA 复性时的起始浓度，*t* 是以 s 为单位的时间。在研究 DNA 序列对复性速率的影响时，将温度、离子强度、核酸片段大小等其他因素均予以固定，以不同时段的复性率取对数后对 *Cot* 作图，可以得到如图 7-5 所示的曲线，用非重复碱基对数目表示核酸分子的复杂度，如 poly（dA）的复杂性为 1，重复的（GATC）$_n$ 组成的 DNA 复杂度为 4，分子长度是 105 bp 的非重复 DNA 的复杂度为 105。原核生物基因组均为非重复序列，故以非重复碱基对表示的复杂度直接与基因组大小成正比，对于真核生物基因组中的非重复片段也是如此。在标准条件下（一般定为 0.18 mol/L 阳离子浓度，400 bp 长的片段）测得的复性率达 0.5 时的 *Cot* 值称为 $Cot_{1/2}$，它与 DNA 序列复杂度成正比。对于原核生物如大肠杆菌的染色体 DNA，该值可代表基因组的大小及基因组中碱基序列的复杂度。真核基因组中因含有许多不同程度的重复序列（repetitive sequence），所得到的 *Cot* 曲线由若干个 S 形曲线叠加而成。总之，复性动力学可用来测定某种生物基因组的大小和特征以及重复序列的拷贝数。

2. 核酸杂交

核酸杂交（nucleic acid hybridization）是一种利用核酸分子的变性和复性的性质，将来源不同的核酸片段按照碱基互补配对规则形成异源双链（heteroduplex），进而对特定目标核酸进行定性或定量分析的技术。异源双链可以在 DNA 与 DNA 之间，也可在 RNA 与 DNA 之间形成。

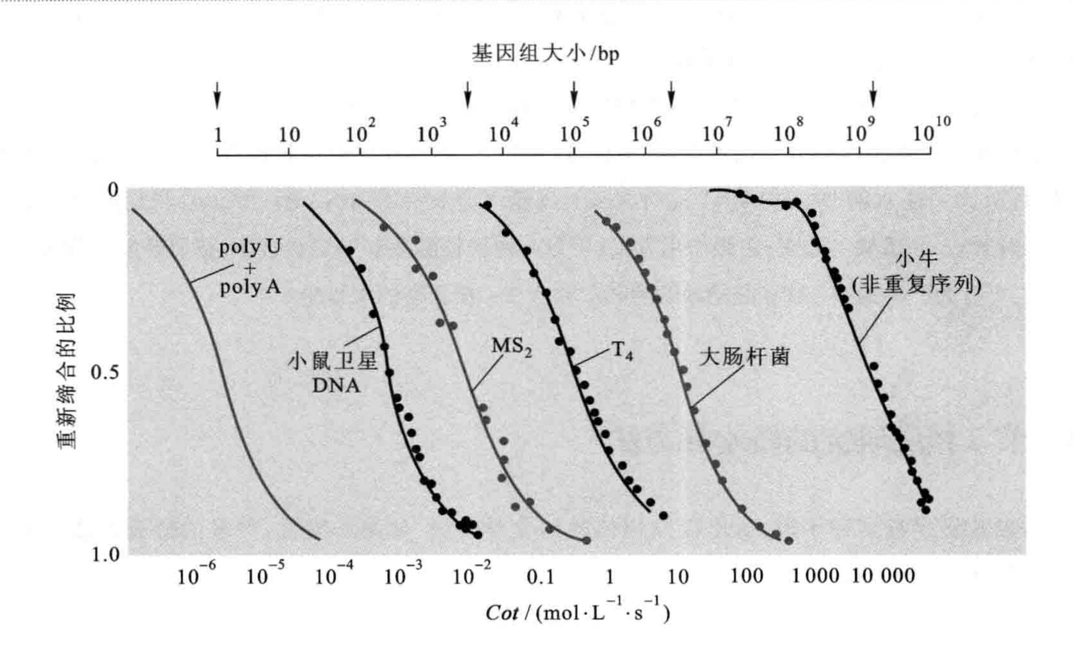


图 7-5 不同 DNA 的复性动力学曲线

框 7-1 生化动态——肽核酸的发现及其应用

肽核酸(peptide nucleic acid,PNA)是 20 世纪 90 年代初人们通过模拟核酸的化学结构,经人工合成得到的 DNA 或 RNA 的类似物。与核酸不同的是,PNA 的主链骨架是由重复排列的 *N*-2- 氨乙基 - 甘氨酸(*N*-2-aminoethyl-glycine)单位经由肽键连接而成,而不是磷酸戊糖骨架。与 DNA 相同的是,PNA 也携带 ATCG 四种碱基,因此也能够按照碱基互补的原则,识别相应的互补碱基序列,但碱基与骨架之间是以亚甲羰键而不是 N- 糖苷键相连。与多肽链相似的是,PNA 也有 N 端与 C 端(图 7-6)。

由于 PNA 没有带高度负电荷的磷酸基团,因此 PNA 与 DNA 之间缺乏电荷相斥的现象,使两者之间的结合强度大于 DNA 与 DNA 之间的结合强度。此外,肽核酸不被目前已知的任何天然核酸酶或蛋白酶所降解,且碱基配对特异性极强,热稳定性高。鉴于以上几点优势,PNA 在很多地方可以代替天然的 DNA,作为一种工具被广泛用于生化、遗传和医学等领域。例如将合成的 PNA 进行标记,即可作为分子杂交技术的探针,对靶核酸实施检测。有一些假说认为,早期地球上的生命形态是以 PNA 为遗传物质,但目前缺乏任何证据。

DNA

PNA

图 7-6 DNA 与 PNA 的结构比较

核酸杂交可以在液相或固相中进行,它已成为核酸研究中一项常规的技术(参看第四十四章“重组 DNA 技术”),像 Southern 印迹、Northern 印迹和 DNA 芯片都是建立在此项技术的基础之上。在医学上,该技术目前已应用于多种遗传性疾病的基因诊断、传染病病原体的检测和恶性肿瘤的基因分析等。

四、核酸的水解

酸、碱和酶均可导致核酸水解。

(1) 酸水解 核酸分子内的糖苷键和磷酸二酯键对酸的敏感性不同:糖苷键 > 磷酸酯键,其中嘌呤糖苷键 > 嘧啶糖苷键。例如,将核酸在 pH 1.6、室温下对水透析,或者在 100℃下、pH 2.8 的溶液中存放 1 h,多数嘌呤碱基即可脱落。核酸的脱嘧啶作用需要在更加剧烈的条件下进行,如使用 98%~100% 甲酸在 175℃下作用 2 h,或者使用三氟乙酸在 155℃下作用 60~80 min,多数嘧啶碱基才会脱落。

(2) 碱水解 RNA 特别是 mRNA 分子内的磷酸二酯键对碱异常敏感。在室温下,0.3~1 mol/L 的 KOH 溶液在约 24 h 可将 RNA 完全水解,并得到 2′- 或 3′- 核苷酸的混合物。DNA 对碱的作用并不敏感,

Quiz5 为什么 RNA 在碱性溶液中比在酸性溶液中更容易水解?

Quiz6 分别将正超螺旋和负超螺旋质粒 DNA 与核酸酶 S1 保温在一起，经过一段时间以后，它们的结构会不会发生什么变化？为什么？

e7-1 常见的水解酶的名称和性质

其抗碱水解的生理意义在于作为遗传物质的 DNA 应更稳定，不易水解。而 RNA（主要是 mRNA）是 DNA 的信使，完成任务后应该迅速降解。

(3) 酶促水解　核酸可受到多种不同酶的作用而发生水解，但不同的酶对底物的专一性、水解的方式和磷酸二酯键的断裂方式是不同的，因此可以按照上述性质对有关的酶进行分类。按照底物特异性，可分为只能水解 DNA 的 DNA 酶（DNase）、只能水解 RNA 的 RNA 酶（RNase）和既能水解 DNA 又能水解 RNA 的磷酸二酯酶；按照作用方式，可分为内切核酸酶和外切核酸酶；按照磷酸二酯键的断裂方式，可分为产物为 5′－核苷酸的水解酶和产物为 3′－核苷酸的水解酶。

第二节　核酸研究的技术和方法

核酸的研究有多种手段，这里仅介绍核酸的化学合成、分离和纯化，更多的研究方法参看第四十四章“重组 DNA 技术”。

一、核酸的化学合成

核酸的化学合成已成为生物化学和分子生物学研究中不可或缺的一项技术，PCR 引物、测序引物、DNA 芯片、基因工程、DNA 探针、核酸适体和干扰 RNA 等都需要使用化学的手段合成长短不一的核酸。此外，合成的寡聚核苷酸还可以应用在疾病的诊断和治疗上。

与生物合成的方法相比，核酸的化学合成具有以下几点好处：不受序列和结构的限制；在合成中可直接引入修饰的碱基、修饰的糖或修饰的核糖－磷酸骨架，从而提高特异性、碱基配对的稳定性，甚至还能产生新的功能。

e7-2 核酸的化学合成

核酸的化学合成与多肽的化学合成在原理上基本相似，如都需要对活性基团进行保护或去保护，对成键基团进行活化，且一般都使用固相合成法等。基本步骤包括：起始原料的制备、偶联反应和去保护。

二、核酸的分离、纯化和定量

（一）核酸的抽取

(1) 两种核蛋白的分离　核酸在细胞内通常以核蛋白的形式存在。其中 RNA 以核糖核蛋白（ribonucleoprotein）、DNA 以脱氧核蛋白的形式存在。借助两种核蛋白在不同盐浓度下溶解度的差别可将二者分开，脱氧核蛋白在 0.14 mol/L NaCl 溶液中的溶解度很低，在 1 mol/L NaCl 溶液中很高，而核糖核蛋白在 0.14 mol/L NaCl 溶液中的溶解度较高。因此，常用 0.14 mol/L 和 1 mol/L NaCl 溶液分别抽取核糖核蛋白和脱氧核蛋白。

(2) 蛋白质的去除　一旦得到核蛋白，就需要将与核酸结合的蛋白质除去。去除蛋白质的方法包括蛋白质酶 K 的消化和酚 / 氯仿的多次抽取。如果抽取 DNA，可先用 RNA 酶消化去除残留的 RNA；如果抽取 RNA，可事先用 DNA 酶尽可能除去残留的 DNA。在酚 / 氯仿抽取中，核酸溶解在上层水相，而蛋白质变性后处于两相的界面。

Quiz7 在用酚 / 氯仿抽取 DNA 的时候，DNA 位于上层还是下层？

(3) 核酸的沉淀　在酚 / 氯仿抽取以后，水相中的核酸可在一定盐浓度下，使用 2.5~3 倍体积的冷无水乙醇进行沉淀。如果纯化的是 RNA，尤其是 mRNA，需要特别小心，务必要采取各种必要的措施来防止 RNA 的降解。

（二）电泳

核酸一般在正常的 pH 下均带负电荷，因此可使用电泳对不同大小的核酸进行分离、鉴定。用于核酸的电泳方法有琼脂糖凝胶电泳和聚丙烯酰胺凝胶电泳。使用最多的是琼脂糖凝胶电泳，而聚丙

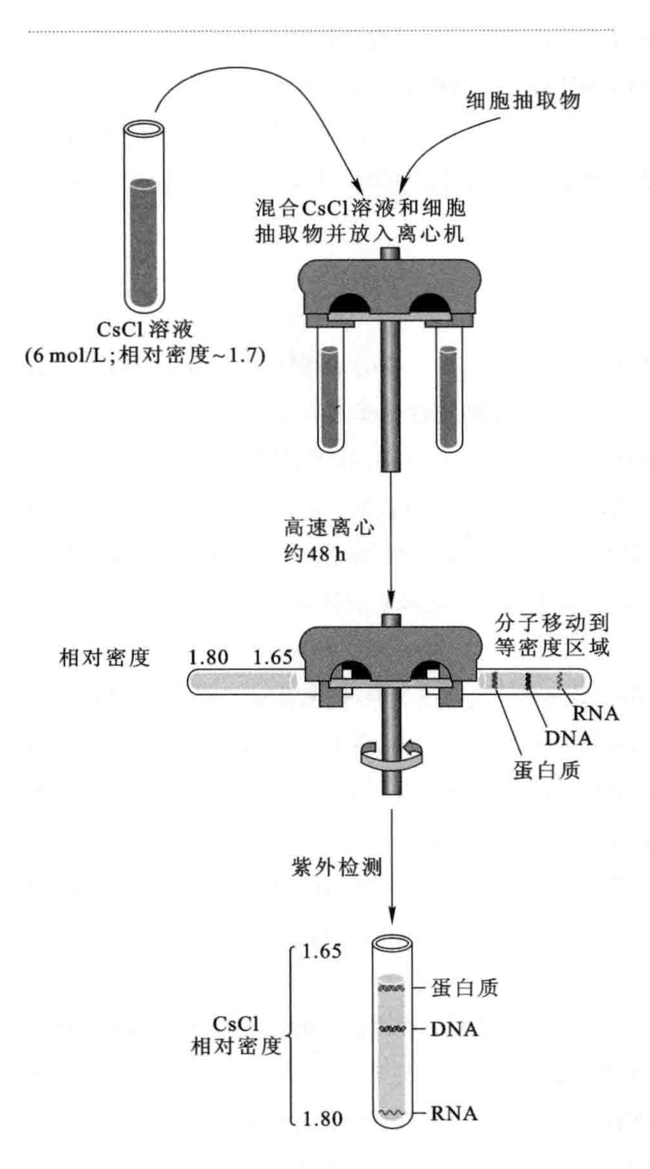


图 7-7 核酸的密度梯度离心分离(Garrett 等,2010)

烯酰胺凝胶电泳一般用于 DNA 序列分析和分离较小的核酸(如 PCR 的引物)。

如果使用琼脂糖凝胶电泳,可使用溴乙锭(ethidium bromide,EB)染色进行检测,因为 EB 可插入到 DNA 双螺旋的碱基对之间,在 UV 照射下发出荧光;如果使用聚丙烯酰胺凝胶电泳,一般用放射自显影或银染等法进行检测。

(三) 离心

离心也是核酸研究中的一项常见技术,它除了可以用来收集沉淀的 DNA,还可以用来进一步纯化核酸,获得高纯度的 DNA。此外,还可以用它来测定一种 DNA 分子中的 GC 含量。图 7-7 显示的是使用 CsCl 密度梯度离心分离 DNA、RNA 和蛋白质的流程。由于这三种大分子具有不同的浮力密度,在 CsCl 梯度溶液中,经过一定时间离心后,它们将处于不同的密度区域。RNA 密度最高,所以位于离心管底,而蛋白质最轻,将位于上方,而 DNA 则处于它们之间的某一位置。

Quiz8 同一种 DNA,变性的和没有变性的在 CsCl 密度梯度离心中所处的位置有无变化? 为什么?

(四) 层析

各种层析蛋白质的方法同样可以用来纯化核酸。如利用阴离子交换层析分离制备核酸,羟基磷灰石分离单链 DNA 和双链 DNA,寡聚 dT 亲和层析分离带有多聚腺苷酸尾巴的真核生物 mRNA。

(五) 核酸纯度检测与核酸定量

核酸纯度检测和核酸定量的最简单方法是紫外分光光度法,通过测定 OD_{260}/OD_{280} 比值来推算。对于 DNA 来说,如果比值大于 1.9,则可视为纯度较高,如果小于 1.9,则可能有蛋白质污染;对于 RNA 来说,如果比值在 1.8~2.0,则可视为纯度较高。

对于纯的 DNA 来说,1 个 OD_{260} 相当于 50 μg/ml 双链 DNA 或 35 μg/ml 单链 DNA;对于纯的 RNA 来说,1 个 OD_{260} 相当于 40 μg/ml RNA。

第三节 核酸一级结构的测定

快速、准确地测定出一种核酸分子的一级结构具有十分重要的意义。在医学上,不仅可以帮助医务工作者迅速确定出流行病暴发或遭遇生物恐怖袭击时的病原体,以拯救许多宝贵的生命,而且还可以帮助医生诊断出与遗传性疾病有关的基因,从而找到合适、有效的治疗方法。

1975 年以前,确定核酸的一级结构要比测定蛋白质的一级结构困难得多,其主要原因是核酸只含有 4 种核苷酸,而蛋白质却有 20 多种氨基酸,显然,能够对核酸进行选择性切割的特异性位点很少,于是识别特定的核苷酸序列就很困难,不确定性很大。同时,大多数核酸所含的核苷酸数目要比多

肽链上的氨基酸数目多得多，这就进一步增加了测序的难度。使核酸序列测定发生革命性变化的因素主要有两个：其一是发现了能够识别特定核苷酸序列并对 DNA 进行特异性切割的限制性内切核酸酶（restriction endonuclease，RE），利用 RE 的这种性质，可以将一个长的核酸分子定向切割成若干可操作的片段；其二是聚丙烯酰胺凝胶电泳技术的发展，使人们能够将大小仅差 1 个核苷酸的核酸片段分开。

一、DNA 一级结构的测定

目前用来测定 DNA 序列的基本方法有两种，一是由 Sanger 发明的末端终止法（chain termination method），另外一种是由 M. Maxam 和 W. Gilbert 发明的碱基特异性化学断裂法（base-specific chemical cleavage method）。这两种方法一般都会用放射性 ^{32}P 对 DNA 进行标记，并使用聚丙烯酰胺凝胶电泳的方法对长度不同的核酸片段进行分离，分离以后还需要使用放射自显影（autoradiography）技术进行观察和分析。目前已普遍使用荧光物质代替放射性同位素对 DNA 进行标记。由于同位素和荧光标记这两种方法的灵敏度都很高，纳克（ng）级的 DNA 样品就能得到很满意的结果。

1. 末端终止法

末端终止法也叫双脱氧法（dideoxy method）。要想理解此方法的原理，需要对 DNA 复制的过程有所了解（详见第三十三章“DNA 复制”）。DNA 复制是在 DNA 聚合酶催化下，以亲代 DNA 的两条母链上的碱基序列为模板，按照碱基互补配对的原则合成新的 DNA 分子的过程。复制需要引物和 4 种 dNTPs，且总是从 5′端向 3′端进行。细胞内复制的引物一般是 RNA，但体外 DNA 复制的引物是人工合成的与模板链互补的一段寡聚脱氧核苷酸。复制开始于引物 3′端的自由羟基，根据 DNA 模板链的序列合成互补的序列，不断地形成新的 3′，5′－磷酸二酯键，使 DNA 链得到延伸，直到一个新的 DNA 分子被完全合成。

Quiz9 使用双脱氧法测序，需要使用高保真的即具有 3′－外切酶活性的 DNA 聚合酶吗？为什么？

实际使用末端终止法测序的时候，通常使用一种经过基因工程改造过的 T7 噬菌体 DNA 聚合酶来催化测序反应，引物是人工合成的寡聚脱氧核苷酸。首先需要进行 4 组平行的测序反应，每组反应均使用相同的模板，相同的引物以及 4 种 dNTPs（有一种要用放射性同位素标记），即 dATP、dGTP、dCTP 和 dTTP，并在每组反应中各加入一种适量的 2′，3′－双脱氧核苷酸（2′，3′-dideoxynucleotide，ddNTP），如 ddATP，使其随机参入到 DNA 链中。由于 ddNTP 缺乏 3′－羟基，因此一旦参入到 DNA 链中，将导致 DNA 合成的末端终止，从而产生相应的 4 组具有特定长度、不同长短的 DNA 片段，而每一组内的 DNA 片段以同样的双脱氧核苷酸结尾。然后将 4 组 DNA 片段再经过聚丙烯酸胺凝胶电泳按链的长短分开，最后经过放射自显影技术，就可以自下而上直接读出被测 DNA 的核苷酸序列（图 7–8）。以图 7–8 为例，直接读出来的序列应该是 5′AGCGTAGC 3′，该序列应与原来的待测链的序列互补，因此，原来的作为模板链的序列就是 5′GCTACGCT 3′。

2. 碱基特异性化学断裂法

碱基特异性化学断裂法可简称为化学断裂法或 Maxam-Gilbert 法，其基本原理是用特殊的化学试剂处理待测的已在末端被放射性同位素（^{32}P）标记的单链 DNA，或者只有一条链的末端被放射性同位素标记的双链 DNA，造成其特定碱基的修饰、脱落和戊糖－磷酸骨架被特异性切割，产生一组长度不同的 DNA 链裂解产物。再用聚丙烯凝胶电泳分离和放射自显影观察，最后可直接读出待测 DNA 片段的核苷酸序列（图 7–9）。

使用的碱基修饰试剂有两类，一类为硫酸二甲酯（dimethylsulphate，DMS），针对 A 或 A+G，另一类为肼，针对 C 或 C+T。使用化学断裂法测序同样需要进行 4 组平行的反应：

（1）G 特异性反应　在碱性条件下，DNA 受 DMS 的作用，其链上的 G 在 N7 位发生甲基化修饰。甲基化的 G 与核糖之间的糖苷键变得不稳定，再经哌啶（piperidine）的作用，嘌呤环被打开并且发生脱落，随后与无 G 的脱氧核糖环相连的磷酸二酯键断裂，结果是每遇到一个 G 就产生两个 DNA 片段，但

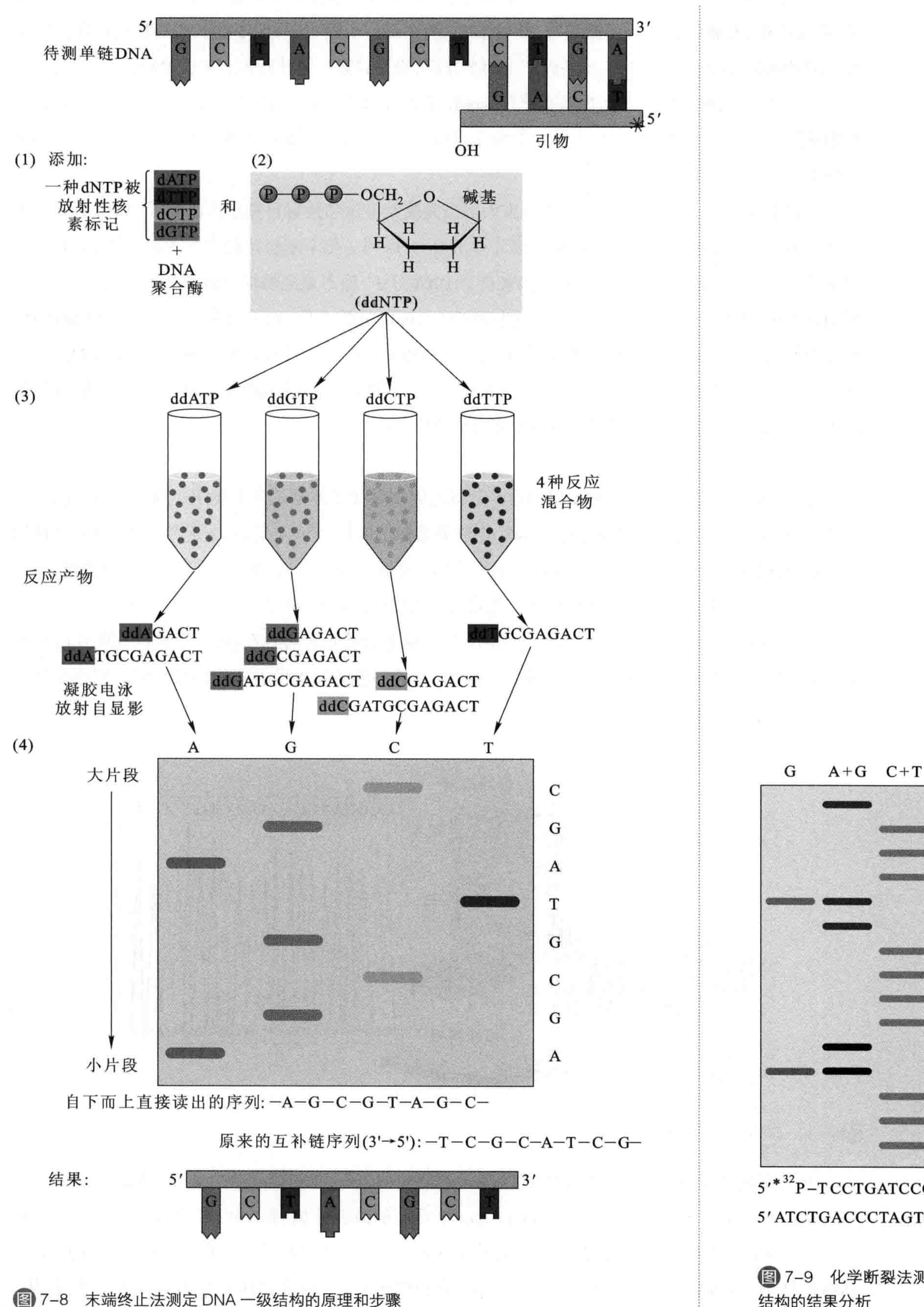


图 7-8 末端终止法测定 DNA 一级结构的原理和步骤

图 7-9 化学断裂法测定 DNA 一级结构的结果分析

只有一个片段带有同位素标记。

(2) 嘌呤碱基特异性的反应 先对 DNA 进行酸处理，然后再加 DMS。这样的条件会导致 DNA 链在 G 的 N7 和 A 的 N3 位都发生甲基化。随后的处理方法同(1)，结果是每遇到一个嘌呤碱基就会产生两个片段，同样只有一个片段有同位素标记。

(3) 嘧啶碱基特异性的反应　首先在肼作用下，DNA 链上嘧啶环发生水解而打开。再受哌啶的作用，嘧啶脱落，裸露的脱氧核糖被修饰，并发生 β- 消去反应。最后，与无嘧啶的脱氧核糖环相连的磷酸二酯键断裂，结果是每逢嘧啶碱基就产生两个片段，也是只有一个片段带有同位素标记。

(4) C 特异性的反应　在高盐浓度下（1 mol/L 或 2 mol/L 的 NaCl），按照(3)的方法处理 DNA，这时 T 受到保护，不会与肼起反应，只有 C 才会发生反应。于是，每逢 C 就产生两个片段，仍然是只有一个片段带有同位素标记。

在以上 4 组反应结束以后，就可进行聚丙烯酰胺凝胶电泳和放射自显影。比较 G、A+G、C+T 和 C 各个泳道，自下而上从自显影 X 光片上就可读出 DNA 序列。化学断裂法较之末端终止法，具有一个明显的优点，那就是测定出来的序列直接来自原 DNA 分子，而不是经酶促合成产生的新拷贝。因此，利用化学断裂法可对人工合成的 DNA 进行测序，也可以分析天然 DNA 原来可能含有的修饰碱基。还可以结合蛋白质保护及修饰干扰实验，测定一个 DNA 分子上一段特殊的碱基序列，如 DNA 酶 I- 足印法测定启动子序列（参看第三十六章“DNA 转录”）。然而，由于末端终止法既简便又快速，其使用的化学试剂的毒性低，它仍然是当今 DNA 序列测定的首选。

3. DNA 序列分析的自动化

目前，基于末端终止法的 DNA 序列自动分析已完全替代了原来的手工测定。这需要将不同荧光标记的 DNA 引物引入到测序反应中，例如，红色荧光标记引物用于 A 反应，蓝色用于 T 反应，绿色用于 G 反应，黄色用于 C 反应。首先按照标准的末端终止法进行测序反应。然后，将 4 组反应混合物合并，并在同一块凝胶的同一个泳道进行电泳。随着电泳的进行，各个寡核苷酸片段在胶上被分开，并按照从小到大的次序依次通过凝胶的底部。受氩激光器发生的激光束的激发，每一个寡核苷酸片段在 5′ 端的荧光标记发出荧光。荧光的颜色被自动检测，不同颜色的荧光代表不同的核苷酸。测序的最后结果可直接打印出来（图 7-10）。

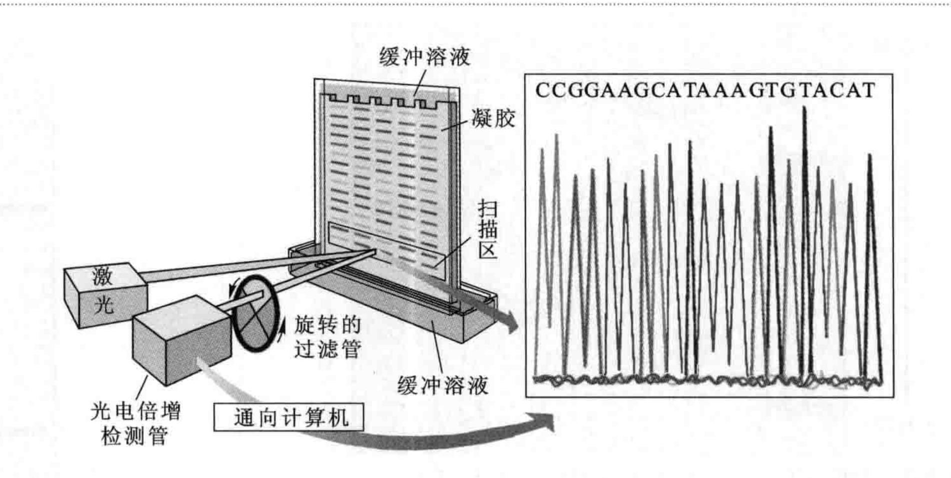


图 7-10　DNA 序列自动分析仪的组成

DNA 序列分析的自动化大大加快了 DNA 序列测定的进程，为科学家测定某一个物种的全基因组序列提供了可能。根据 http://genomesonline.org 主页上的最新数据，到 2012 年 7 月 1 日，共有 3 383 种生物的全基因组序列已完成测定，其中古菌 155 种，细菌 3 053 种，真核生物 175 种。它们包括：流感嗜血杆菌、大肠杆菌、酿酒酵母（120 万 bp）、秀丽线虫（17 500 个基因）、拟南芥（2.6 万个基因）、果蝇（13 601 个基因）、小鼠（近 3 万个基因）、人（2 万多个基因）、水稻（5 万多个基因）、杨树（4.5 万多个基因）、高粱（3 万多个基因）、狗、河豚和蚊子等。

4. DNA 序列测定的新方法

除了上述 2 种常见测序方法以外，近几年来，科学家还发明了一些新的测序方法，例如焦磷酸测序法（pyrosequencing）和单分子测序法（single-molecule sequencing）。其中焦磷酸测序技术无需电泳，

DNA 片段也无需荧光标记，但可以在芯片上进行高通量分析。建立在这两种测序方法基础之上的高通量测序技术堪称测序技术发展历程的一个里程碑，该技术可以对数百万个 DNA 分子同时进行测序，操作极为简便，大大节约了成本和时间。这使得对一个物种基因组和转录组进行细致全面的分析成为可能，因此也称其为深度测序（deep sequencing）。

e7-3 焦磷酸测序和深度测序

二、RNA 一级结构的测定

目前用来测定 RNA 一级结构的方法主要有：①先使用逆转录酶将待测的 RNA 逆转录成 cDNA，然后再使用 Sanger 的末端终止法进行测定；②先用化学和 / 或酶学方法对放射性同位素标记的 RNA 进行部分消化，再进行聚丙烯酰胺电泳分析；③质谱法。

框 7-2　生化趣事——年度分子和年度突破

受 *Time* 杂志每年评选的年度人物（man of the year）的启发，*Science* 在 1989 年开始评选年度分子（molecule of the year），以让人们认知科学界的重大发现。但从 1996 年开始，科学杂志将年度分子改为年度突破（breakthrough of the year）。

今年的年度突破是发现口服抗艾药物可防止 HIV 在异性之间的转播。第一个被评为年度分子的是一种来自水生嗜热菌的 *Taq* DNA 聚合酶，它早已在 PCR 反应中得到广泛的应用。到 2011 年底，已经有 21 个年度分子或年度突破上榜（表 7-1），其中与生化有关的有 12 个，如 2000 年的年度突破为全基因组测序。

► 表 7-1　科学杂志的年度分子与年度突变

年度	年度分子 / 年度突破	年度	年度分子 / 年度突破
1989	*Taq* DNA 聚合酶	2001	纳米电路
1990	人造钻石	2002	小 RNA（small RNAs）
1991	富勒烯（buckminsterfullerene）	2003	暗能量
1992	一氧化氮（NO）	2004	勇气号火星车
1993	p53 蛋白	2005	进化
1994	DNA 修复酶	2006	庞卡赫猜想的证据
1995	玻色 - 爱因斯坦凝聚态	2007	人类遗传变异
1996	艾滋病毒（HIV）	2008	细胞再编程（cellular reprogramming）
1997	多莉羊（Dolly the sheep）	2009	拉密达猿人（*Ardipithecus ramidus*）
1998	暗物质	2010	第一个量子机器
1999	干细胞	2011	口服抗艾药物可防止 HIV 在异性之间的转播
2000	全基因组测序		

推荐网址：

e7-4 本章小结填填看

1. http://genomesonline.org（基因组在线网站，实时更新最新的基因组数据）
2. http://www.dnalc.org/resources/animations/sangerseq.html（内有动画显示双脱氧序列测定过程，还有其他生物化学和分子生物学动画资源）

参考文献：

1. Sanger F, et al. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA*, 74(12): 5 463–5 467, 1977.
2. Maxam A M, Gilbert W. A new method for sequencing DNA. *Proc Natl Acad Sci USA*, 74(2): 560–564, 1977.

第八章 酶学概论

一个独立的生命系统包含的化学反应虽然有成百上千种，但几乎所有的反应都是在高效、专一的催化剂的催化下进行的。酶(enzyme)就是由细胞合成的，在机体内行使催化功能的生物催化剂(biocatalyst)，其化学本质主要是蛋白质，极少数是RNA。可以说，没有酶，生命是不可能存在的。正因为酶是如此奇妙和重要，所以细胞内的任何一种酶的缺失、突变、表达不足、表达过量或活性异常都有可能对细胞的生命活动带来不利的影响，有时这种影响是灾难性的。实际上，很多遗传性疾病是由酶基因的突变引起的。同时，许多治疗疾病的药物也是以特定的酶作为靶点。

本章将主要介绍酶的一般性质，包括酶的化学本质、酶的催化性质和酶的分类与命名。

第一节 酶的化学本质

早在几千年前，古人就无意识地利用了酶的催化作用，例如，利用一些特殊的微生物来治疗疾病、制造食品和饮料。然而，真正认识到酶的存在和作用，是始于19世纪西方国家对酿酒发酵过程进行的大量研究。1833年，法国科学家Anselme Payen和Jean Persoz从麦芽的水抽提物中，用酒精沉淀得到一种对热不稳定的活性物质，可以促进淀粉水解成可溶性的糖，他们把这种物质叫转化酶(diatase)，后来又更名为淀粉酶(amylase)，它也是公认的第一种被发现的酶。1878年，Kunne首先把这种物质称为酶，其词根来自希腊文，意思是“在酵母中”。1834年，德国科学家Theodor Schwann从胃襞中得到第一种动物来源的酶，即胃蛋白酶。1898年，Pierre Émile Duclaux建议在所有酶的名称加上后缀“ase”。

在意识到酶的存在及其对生命的重要性以后，人们开始探究它的化学本质，但要取得突破并非易事。当时以R. M. Willstarter为代表的科学家认为，酶既不是蛋白质，也不是糖或脂肪，只是一种吸附在蛋白质上的活性物质。20世纪30年代，James Batcheller Sumner和John Howard Northrop分别得到脲酶和胃蛋白酶的结晶，并用实验证明了酶是蛋白质。那么，酶都是蛋白质吗？

对上述问题持肯定回答，即“酶都是蛋白质”这个观点曾经长期统治学术界，几乎没有人怀疑过。然而，在1982年，Thomas Cech等发现，四膜虫26S rRNA前体具有自我剪接功能，并于1986年证明其内含子L-19 IVS具有多种催化功能。1984年，Sidney Altman等人发现大肠杆菌核糖核酸酶P的核酸组分——M1 RNA具有酶的活性。Cech和Altaman的发现震惊了全世界，也从此推翻了“酶都是蛋白质”的传统观念。现在一般认为，自然界绝大多数酶是蛋白质，仅有少数为RNA。具有催化活性的RNA称为核酶。核糖体也被证明是一种核酶，蛋白质合成最重要的一步反应，即肽键的形成就是由它催化的。这些事实表明核酶是普遍存在的。并且现在还可设计并合成出一系列非自然界存在的核酶(参看第十一章“核酶的结构与功能”)。

酶蛋白质分子根据化学组成可分为单纯蛋白质和缀合蛋白质，前者称为单纯酶(simple enzyme)，后者称为缀合酶(conjugated enzyme)或结合酶。在缀合酶分子上，除包括由氨基酸残基组成的多肽链以外，还包括与肽链结合的某些对热稳定的有机小分子或金属离子。这些非氨基酸成分统称为辅助因子(cofactor)。丧失辅助因子的酶称为脱辅酶(apoenzyme)，与辅助因子结合在一起的酶称为全酶(holoenzyme)。辅助因子包括辅酶(coenzyme)、辅基(prosthetic group)和金属离子三类。辅酶专指那些

与脱辅酶结合松散、使用透析或超滤等温和的方法就能去除的有机小分子，如辅酶Ⅰ和辅酶Ⅱ。辅基专指那些与脱辅酶结合紧密（有时甚至以共价键结合）、使用透析或超滤的方法难以去除的有机小分子，比如琥珀酸脱氢酶中的FAD。含有紧密结合的金属离子的酶通常称为金属酶。最常见的充当辅助因子的金属离子有铜、镁、锌和锰。

酶RNA分子中，少数仅由RNA组成以外，绝大多数还含有金属离子或/和蛋白质。由于自然界绝大多数酶为蛋白质，因此下文中有关酶学的内容主要是围绕化学本质为蛋白质的酶展开的，至于核酶的内容多数集中在第十一章“核酶的结构与功能”做介绍。

根据酶蛋白本身结构的特征，酶又可分为单体酶、寡聚酶和多酶复合物，它们的性质比较参看表8–1。

► 表8–1　单体酶、寡聚酶和多酶复合物的性质比较

名称	组成	M_r	实例
单体酶	只有1条肽链	13 000~35 000	绝大多数水解酶
寡聚酶	≥2个亚基，以次级键结合	35 000至百万	许多调节酶
多酶复合物	≥2个功能相关的酶嵌合而成	约几百万	丙酮酸脱氢酶系

单体酶中有一类酶，虽然只由一条肽链组成，但同时具有多种不同的酶活性，这类单体酶称为多功能酶（multifunctional enzyme），如大肠杆菌DNA聚合酶Ⅰ就同时具有DNA聚合酶、3′–外切核酸酶和5′–外切核酸酶的活性（参看第三十三章“DNA复制”）。

第二节　酶的催化性质

酶作为生物催化剂，与非酶催化剂一样，只能催化热力学允许的反应，反应完成后本身不被消耗或改变，即可以重复使用。它对正反应和逆反应的催化作用相同，不改变平衡常数，只加快到达平衡的速率或缩短到达平衡的时间。受酶催化的化学反应称为酶促反应，其中的反应物称为底物（substrate）。但作为生物催化剂，酶还具有一些特有的性质，它们包括：

1. 高效性

酶催化效率之高是无与伦比的。酶促反应一般要比无催化剂的反应速率高10^6~10^{12}倍，有些反应更高。例如，由乳清苷–5′–磷酸脱羧酶催化的反应要比无催化剂的反应快10^{17}倍！酶促反应比非酶催化剂催化的反应至少也要高几个数量级。酶催化的高效率与其能够大幅度降低反应的活化能（activation energy）有关，详情可以参看表8–2中有关过氧化氢酶的数据。

► 表8–2　相对反应速率与相对活化能之间的关系

反应条件	相对活化能	相对反应速率
无催化剂	18 000	10^{-7}
Fe催化剂	12 000	46
过氧化氢酶	2 000	4×10^6

2. 酶在活性中心与底物结合

酶的活性中心（active site）也被称为活性部位，是指酶分子上直接与底物结合，并与催化作用直接相关的区域。如果酶是缀合酶，活性中心还包括与辅助因子结合的区域。如果一种酶是多功能酶，就会有多个活性中心。

活性中心是由结合基团和催化基团组成。前者负责与底物结合，决定酶的专一性，后者参与催化，

Quiz1 哪些氨基酸残基的侧链不可能参与催化？

负责底物旧键的断裂和产物新键的形成，决定酶的催化能力。但也有某些基团可能兼有两项功能。

活性中心一般具有以下特征：

(1) 活性中心是一个三维实体，通常由在一级结构上并不相邻的氨基酸残基组成　构成活性中心的氨基酸残基和辅助因子的所有原子精确有序地排列在一起，其独特的三维实体结构是整个蛋白质正确折叠后的必然产物。例如，构成溶菌酶活性中心的基团包括 1、35、52、62、63、101、108 和 129 位的氨基酸残基，当溶菌酶正确折叠后，这些氨基酸残基自然排列在一起，共同组成活性中心。

(2) 活性中心只占酶总体积的一小部分(1%~2%)　酶分子上的大多数氨基酸残基并不与底物接触，但它们作为结构支架，有助于稳定活性中心的三维结构。

(3) 活性中心为酶分子表面的一个裂缝(cleft)、空隙(crevice)或口袋(pocket)，中心内多为疏水氨基酸残基，也有少量亲水氨基酸残基，以便底物的结合和催化的进行。

活性中心如此设计的目的是把降低酶催化活性的溶剂(水)排除在外，防止副反应的发生。实际上，除非作为底物，水分子通常被排除在活性中心之外，而底物分子在进入活性中心之前则需要去溶剂化。

Quiz2 为什么绝大多数酶的活性中心含有 His？

到目前为止，在已被研究的酶中，有 65% 以上的酶活性中心含有 His、Cys、Asp、Arg 或 Glu，这 5 种氨基酸残基出现的频率是 His>Cys>Asp>Arg>Glu。

(4) 与底物结合为多重次级键，包括氢键、疏水键、盐键和范德华力。

(5) 底物结合的特异性在一定程度上取决于活性中心和底物之间在结构上的互补性，但活性中心与反应过渡态的互补性要好于与底物的互补性。

(6) 活性中心的构象不是固定不变的，而是具有一定的柔性。

3. 高度的专一性

酶的专一性(specificity)是指酶对参与反应的底物有严格的选择性，即一种酶仅能作用于一种底物或一类分子结构相似的底物，使其发生某种特定类型的化学反应，并产生特定的产物。

不同酶的专一性是不一样的。有的酶只能作用于一种底物，如碳酸酐酶只能作用于碳酸，葡糖激酶只能作用于葡萄糖。有些酶的专一性就比较低，例如己糖激酶能作用于葡萄糖和甘露糖，再如胰凝乳蛋白酶不仅能够水解蛋白质，还能水解酯。

专一性一般有 4 种类型：

(1) 绝对专一性(absolute specificity)　是指一种酶仅催化一个特定的反应。例如，脲酶只能催化尿素的水解反应。甲基尿素(methylurea)与尿素的结构非常相似，但脲酶对其无任何作用。

(2) 基团专一性(group specificity)　是指一种酶只作用于含有特定官能团(如磷酸基团、氨基和甲基等)的分子。如磷酸酶只水解特定底物分子上的磷酸基团。

(3) 键专一性(linkage specificity)　是指一种酶只作用于含有特定化学键的分子，而不管底物分子其他部分的结构。如二肽酶专门识别二肽中的肽键，而不管构成这个肽键的两个氨基酸残基是哪种。

(4) 立体专一性(stereospecificity)　是指酶对具有立体异构体的底物只作用于其中的一种，而对另外一种无效的性质，进一步可分为旋光异构专一性和几何异构专一性两类。其中旋光异构专一性是指当底物具有旋光异构体时，酶只能作用于其中的一种。如 D－氨基酸氧化酶只作用于 D－氨基酸，而对 L－氨基酸不起任何作用。同样，乳酸脱氢酶只作用于 L－乳酸。再如氨酰－tRNA 合成酶只结合 L－氨基酸，而不结合 D－氨基酸，正因为如此，蛋白质分子中没有 D－氨基酸。几何异构专一性是指酶对几何异构体的专一性。例如，琥珀酸脱氢酶只能催化琥珀酸脱氢而生成反－丁烯二酸，或者催化使反－丁烯二酸得到氢还原成琥珀酸的逆反应，但它并不能催化顺－丁烯二酸的生成或加氢。

酶的立体异构专一性还表现在许多酶能够区分假手性 C 上的两个等同的基团，并且只催化其中的一个，而不催化另一个。例如，一端由 ^{14}C 标记的甘油，在甘油激酶(glycerol kinase)的催化下能与

ATP 反应，但仅产生一种标记产物 1– 磷酸甘油。甘油分子中的两个—CH_2OH 基团在我们看来是完全相同的，可是酶却能区分它们。另外，用氘(D)标记的方法发现，在脱氢酶的催化下，底物和 NAD^+ 之间发生的氢的转移也有着严格的立体异构专一性，这种专一性表现在对尼克酰胺环中 C4 上的氢有选择性。如酵母乙醇脱氢酶(yeast alcohol dehydrogenase，YADH)在催化时，辅酶的尼克酰胺环 C4 上只有一侧是可以加氢或脱氢的，另一侧则不被作用(图 8–1)。

$$CH_3CHO + \text{NADH(4-D, H)} + H^+ \xrightleftharpoons{YADH} H\text{—}C(OH)(CH_3)\text{—}D + NAD^+ \qquad CH_3CDO + NADH + H^+ \xrightleftharpoons{YADH} D\text{—}C(OH)(CH_3)\text{—}H + NAD^+$$

图 8–1　A 型脱氢酶催化的脱氢反应

酵母乙醇脱氢酶的这种专一性被定为 A 型，凡与酵母乙醇脱氢酶具有相同专一性的酶都称为 A 型脱氢酶，比如苹果酸脱氢酶、异柠檬酸脱氢酶以及乳酸脱氢酶等，否则就称为 B 型脱氢酶，如谷氨酸脱氢酶和 α– 甘油磷酸脱氢酶等。

框 8–1　生化趣事——催化动力、生化武器与投弹手甲虫

投弹手甲虫(bombardier beetle)是天生的生化武器专家，具有一种精妙的防身系统。在它腹部的一侧含有一个特制的燃烧室，里面有两个小囊，分别贮备有氢醌(对苯二酚)和过氧化氢。当捕食者靠近它的时候，两种物质才混合在一起，同时燃烧室壁细胞会立刻分泌出两种酶——过氧化氢酶和过氧化物酶。过氧化氢酶催化过氧化氢分解产生氧气，过氧化物酶则催化氢醌的氧化产生醌(图 8–2)。

反应式分别为：$2H_2O_2 \longrightarrow 2H_2O + O_2$；对苯二酚 + $H_2O_2 \longrightarrow$ 醌 + H_2O

这两个酶催化的威力极强，可确保反应在极短的时间内发生，而导致爆炸。

反应产物有 3 种：热，产生接近 100℃的高温；气体氧气，因热而膨胀可作为推进剂；醌，用来攻击和吓唬靠近它的捕食者(如蚂蚁、青蛙和蜘蛛)。

甲壳虫可将其腹部的燃烧室对准捕食者，然后以脉冲的方式高速喷射灼热难闻的热醌，攻击和吓退靠近它的捕食者。在 2005 年 7 月，恐怖分子企图在伦敦地铁引发爆炸，他们使用过氧化氢和面粉的混合物制造爆炸物，只是当时因为雷管故障爆炸才没有获得成功。显然，他们制造爆炸物的灵感不大可能来自有关投弹手甲虫的“酶弹”知识。要知道，投弹手甲虫在利用催化威力无穷的酶来引爆其自制炸弹的时候，从来都不会失败。



图 8–2　投弹手甲虫及其“酶弹”结构示意图

Quiz3　有人将投弹手甲虫的酶弹结构视为不利于进化论的证据。你认为是为什么？

酶的立体专一性在实践中很有意义，例如某些药物只有某一种构型有生理效用，另外一种构型无效甚至有害，而有机合成的药物只能是消旋产物，但用酶来催化可进行不对称合成。

尽管酶表现出高度的专一性，但某些辅助因子可以被多种不同的脱辅酶使用，例如 NAD^+ 为很多脱氢酶的氢受体。

有 3 个模型可以解释酶作用的专一性：

(1) “锁与钥匙”模型(lock and key model)　该模型早在 1894 年就由 Emil Fischer 提出。在该模型中，酶和底物分别被比作锁和钥匙，锁眼则比作为活性中心，他认为活性中心的构象是固定不变的，底物的结构(形状、大小、电荷的分布、氢键供体或受体以及疏水补丁等)必须与它的结构非常吻合才能结合(图 8–3)。该模型有很多缺陷，因此基本上已被淘汰。例如，它不能解释酶的活性中心为何也能与产物结合并催化逆反应。此外，该模型对解释酶的催化机理也没有任何帮助(锁对钥匙没有任何作用)。

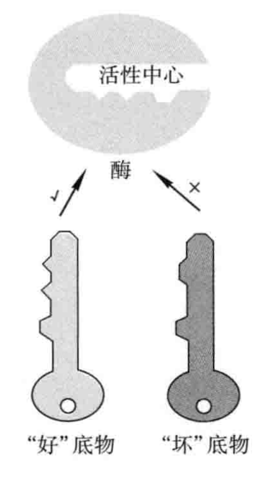


图 8-3 酶与底物结合的"锁与钥匙"模型

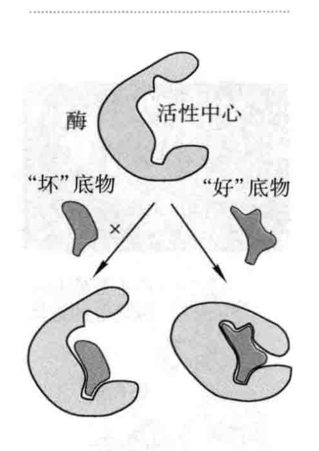


图 8-4 酶与底物结合的"诱导契合"模型

Quiz4 如果用多一个C原子的醛糖代替葡萄糖，结果会是如何？

(2)"诱导契合"模型(induced fit model) 该模型由 Daniel E. Koshland 在 1958 年提出。其主要内容是(图 8-4)：酶活性中心不是僵硬不变的结构，而是具有一定的柔性。酶在与底物结合前后构象是不同的，当底物与酶接近时，酶受到底物分子的诱导，其构象发生变化。有时，底物的构象也会发生变化。这种构象的变化，不仅使得酶能更好地结合底物(有点像戴手套时手套在手的"诱导"下所发生的变化)，还能使活性中心的催化基团处于合适的位置，而能更好地行使催化作用。"好"底物与"坏"底物的差别就在于，前者与酶结合能诱导酶的构象发生有利于催化的变化，而后者不行！

己糖激酶与底物的结合以及随后的催化过程被公认为是最能说明酶"诱导契合"模型的一例(图 8-5)。该酶催化的反应是将 ATP 的一个磷酸基团转移给葡萄糖分子的 6 号位羟基。然而，水和葡萄糖都具有羟基，两者都可以进入己糖激酶的活性中心，但己糖激酶催化磷酸基团从 ATP 转移到葡萄糖分子的效率是转移给水分子的 10^5 倍！对此现象，用"锁和钥匙"模型是不能解释的，但"诱导契合"模型却很容易。因为只有葡萄糖分子进入酶的活性中心以后，才能诱导活性中心从开放的裂缝变成闭合的裂缝。这种构象的变化"赶走"了水分子，拉近 ATP 与葡萄糖 6 号位羟基之间的距离，使活性中心的催化基团处于合适的位置而能更好地行使催化功能，从而保证磷酸基团只会转移到葡萄糖的 6 号位羟基上。因此，从某种意义上来看，葡萄糖通过诱导酶构象的变化而间接地参与了催化反应。

对来源于酵母细胞的己糖激酶所进行的 X 射线衍射实验表明，整个酶分子由两个相对独立的叶(lobe)组成，活性中心为两叶之间的裂缝。当葡萄糖与酶的活性中心结合以后，酶的构象发生了剧烈的变化：作为底物的葡萄糖诱导两叶相向移动，每一叶大约旋转 10°，整个多肽链骨架移动了约 0.8 nm，构成活性中心的裂缝因此而闭合(仿佛河蚌遇到刺激以后两壳闭合)。裂缝的闭合对酶的催化十分重要，原因是：首先它为底物创造了更为疏水的环境，整个葡萄糖分子除了 6 号位的羟基以外都被疏水氨基酸残基的侧链包围，这非常有利于 ATP 的转移。其次，它"赶走"了本来占据在活性中心的水分子，这就防止了酶分子将 ATP 的 γ- 磷酸基团误交给水分子而导致 ATP 发生水解。

有人做过这样的实验：使用少一个 C 原子的木糖代替葡萄糖，结果发现己糖激酶能够催化 ATP 水解。在这种情况下，ATP 之所以发生水解，很可能是因为体积较小的木糖进入活性中心以后，无法赶走"水分子"，而诱使酶"上当受骗"将 ATP 的 γ- 磷酸基团转移给留在活性中心的水分子上的羟基。

上述底物诱导的裂缝闭合(substrate-induced cleft closing)现象不仅仅在己糖激酶上被实现，实际上在其他几种参与糖酵解的激酶(包括磷酸果糖激酶、磷酸甘油酸激酶和丙酮酸激酶)分子上也能够观测到，说明这已成为各种激酶的共同特征。

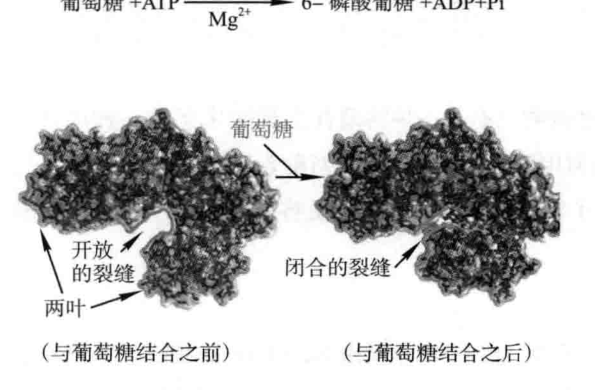
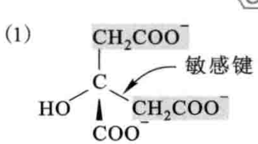


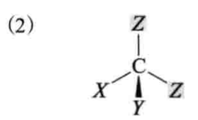
图 8-5 己糖激酶的"诱导契合"

符合"诱导契合"模型的另一例是脂肪酶。该酶催化的反应是脂肪的水解，其底物是脂肪。这种酶对水溶性底物几乎没有任何作用，只能水解位于水相和油相界面的底物。这是因为酶在水相里，活性中心被它的一段螺旋遮盖住，然而到了两相交界处，受到界面的诱导，酶的构象发生变化，活性中心得以外露，于是便有了活性。

(3)"三点附着"模型(three-attachment model) 对于酶为什么能够区分一对对映异构体，或者一个假手性 C 上两个相同的基团(图 8-6 中的 Z)，需要用酶与底物的"三点附着"模型进行解释。该模型认为，底物在活性中心的结合有 3 个结合点，只有当在这 3 个结合点都匹配的时候，酶才会催化相应的反应。一对对映异构体底物虽然基团相同，但空间

排列不同，这就可能出现其中一种与酶结合的时候，无法保证 3 点都互补匹配，酶也就不能作用于它。图 8-6(1) 图显示的是顺乌头酸酶的底物——柠檬酸的立体结构，其中的两个—CH_2—COOH 基团在我们眼里看起来没有什么两样，但在“酶眼”里，它们却是不同的。正因为如此，顺乌头酸酶在催化反应的时候，只会将羟基催化转移给来自草酰乙酸的—CH_2—COOH，而绝对不会转移给来自乙酰 -CoA 的—CH_2—COOH(反应式参看第二十三章“三羧酸循环”)。





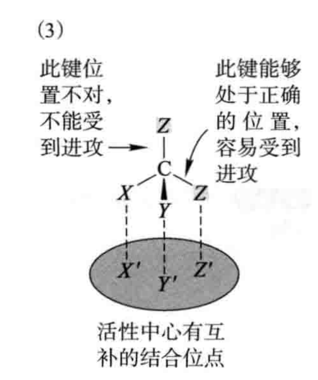


图 8-6　酶与底物结合的“三点附着”模型

“三点附着”模型不但可用来解释酶的立体专一性，而且可以解释其他非酶蛋白质作用的立体专一性。例如，脑细胞中的一种氨基酸受体只能与 L- 氨基酸结合，对 D- 氨基酸没有作用。

4. 反应条件温和

除了发生在一些生活在极端环境下的微生物体内的反应以外，绝大多数酶促反应的条件都十分温和：温度通常是在 37℃左右，101.3 kPa 下，pH 接近 7。

5. 对反应条件敏感，容易失活

与一般的化学催化剂相比，酶对反应条件极为敏感，这与酶的化学本质有关。每一种酶都有最佳的反应条件，如最适 pH 和最适温度等。偏离最佳条件会影响到它的活性，而极端的 pH、特定的抑制剂、过高的温度和压强等因素都会导致酶活性的丧失。

6. 受到调控

酶的活性，特别是一条代谢途径中的限速酶的活性，是受到严格调控的，调控的手段也是各种各样(详见第十二章“酶活性的调节”)。

7. 许多酶的活性还需要辅助因子的存在

作为辅助因子的可能是金属离子，也可能是辅酶或者辅基。作为辅酶或者辅基的多为维生素或其衍生物。有的辅助因子作为第二底物参与反应，如辅酶 Ⅰ 和辅酶 Ⅱ 作为电子和氢的受体或供体参与氧化还原反应，有的也可以直接参与催化，完成氨基酸侧链基团不能完成的催化任务，例如金属离子、焦磷酸硫胺素和磷酸吡哆醛。

酶除了催化特定的化学反应以外，还有以下功能：①控制一个反应何时何地进行；②调节反应的速率；③防止副反应的发生；④优化特殊条件下的反应。

第三节　酶的分类和命名

到目前为止已有数千种的酶被纯化，几百种酶被结晶，随着酶成员的不断扩充，需要对其进行科学的分类和命名。

一、酶的分类

酶的分类是按照国际生物化学和分子生物学命名委员会(Nomenclature Committee of the International Union of Biochemistry and Molecular Biology，NC-IUBMB)的建议，根据反应的性质，将其分为 6 大类，参见表 8-3。

e8-1 六类酶的记忆秘诀

在每一大类酶中，又可根据不同的标准，分为几个亚类。每一个亚类再分为几个亚亚类。每一亚亚类中都含有特定的数字编号，作为一个酶在亚亚类中的顺序号。每一大类、亚类和亚亚类都用具体的数字表示。例如，乳酸脱氢酶编号为 EC 1.1.1.27，其中 EC 表示酶学委员会，4 个数字分别表示此酶属于第一大类(氧化还原酶)、此大类中的第一亚类(氧化基团为 CHOH)、此亚类中的第一亚亚类(NAD^+ 为 H 的受体)和在此亚亚类中的顺序号。

Quiz5 是不是所有酶催化的反应都涉及共价键的断裂?

► 表 8-3　酶的分类及其实例

类别	反应性质	实例
氧化还原酶(oxidoreductase) 包括:脱氢酶,氧化酶,还原酶,过氧化物酶,过氧化氢酶,加氧酶,羟化酶	电子转移	乙醇脱氢酶
转移酶(transferase) 包括:转醛酶和转酮酶,脂酰基、甲基、糖基和磷酸基转移酶,激酶,磷酸变位酶	分子间基团转移	蛋白质激酶 A
水解酶(hydrolase) 包括:酯酶,糖苷酶,肽酶,磷酸酶,硫酯酶,磷脂酶,酰胺酶,脱氨酶,核酸酶	通过加水导致键的断裂	脂肪酶
裂合酶(lyase) 包括:脱羧酶,醛缩酶,水合酶,脱水合酶,合酶,裂解酶	消除反应,产生双键	碳酸酐酶
异构酶(isomerase) 包括:消旋酶,差向异构酶,异构酶,变位酶	分子内的重排	磷酸己糖异构酶
连接酶(ligase) 包括:合成酶和羧化酶	水解 ATP 与分子之间的连接偶联	DNA 连接酶

Quiz6 ▶ 如何区分合成酶和合酶?

二、酶的命名

根据 NC-IUBMB 的建议,每一个酶都被给予两个名称,一个为系统名,一个为惯用名。系统名要求能确切地反映底物的化学本质以及酶的催化性质,因此它由底物名称和反应类型两个部分组成。如果一个酶促反应的底物不止一种,那需要将所有的底物都注明,中间用“:”隔开。例如,乳酸脱氢酶的系统名应该是乳酸:NAD^+ 脱氢酶。

惯用名也需要能反映底物名称和反应性质,但不需要非常准确,一般采用底物加反应类型来命名,比如蛋白水解酶、乳酸脱氢酶、磷酸己糖异构酶等。绝大多数酶的惯用名的英文后缀为“ase”,少数例外,比如胃蛋白酶、胰蛋白酶和肾素的英文名称分别是 pepsin、trypsin 和 rennin。由于惯用名使用起来比较简便,人们更喜欢用它来称呼一种酶。对于水解酶,人们还习惯省去反应的类型,直接用底物来表示,如蛋白酶、核酸酶、脂肪酶、淀粉酶和 ATP 酶就分别表示水解蛋白质、核酸、脂肪、淀粉和 ATP 的水解酶。此外,有时也会在底物名称前冠以酶的来源,如血清谷丙转氨酶和胰蛋白酶。

惯用名简单,应用历史长,但由于缺乏系统性,有时难免出现“一酶数名”或“一名数酶”的现象。

科学故事——酶化学本质的确定

在阐明酶的化学本质过程中,美国科学家 James Batcheller Sumner 功不可没。Sumner 于 1887 年生在美国马萨诸塞州,爱好物理和化学,业余时间喜欢打猎。17 岁外出打猎时,被同伴误伤左臂,不得不截去左前臂,因此他必须学会用右手做事。同时,为了增强体质,他并没有放弃体育活动。1906 年,Sumner 进哈佛医学院专攻化学,1910 年毕业。1912 年,到哈佛医学院随 Otto Folin 教授学习化学,Folin 认为独臂人很难在化学方面获得成功,劝 Sumner 改学法律。Sumner 仍坚持己见,并在 1913 年获得硕士学位,1914 年获博士学位。随后到纽约州康奈尔医学院任教。

1917 年他决定分离酶,并选择脲酶(urease)作为分离对象,但起初并不成功。他在康奈尔医学院打算分离脲酶之前,曾用从大豆中制备的脲酶测定肌肉、血和尿中的尿素含量。1916 年,有人发现南美刀豆中脲酶的含量比大豆多 16 倍。Sumner 认为如此超常量的脲酶是可以用化学方法分离和鉴定的。Sumner 选择富含脲酶的刀豆提取脲酶是他成功的第一步。Sumner 在诺贝尔奖金颁奖仪式上说:“我之所以成功是因为幸运地选择了脲酶。”1917 年,他开始从刀豆中分离和纯化脲酶。1921 年,在他得到美国－比利时合作基金,决定到布鲁塞尔与写过几本有关酶的书籍的 Jean Effront 一起工作。然而,Effront 认为他分离脲酶的想法荒唐可笑,因此计划最终泡汤。回到美国以后,Sumner 继续他的纯化脲酶的研究。1922 年的

一天，他没有使用以往用水、甘油和乙醇提取脲酶的方法，而改用 30% 丙酮。当他取出一滴丙酮抽提液放在显微镜下观察时，发现液体中长出许多小晶体。离心收集这些晶体后，发现它有很高的脲酶活性，分离后的脲酶纯度一下子增加了 700~1 400 倍，这是其他纯化方法难以比拟的。最终 Sumner 分离出了脲酶，此后他又做了一系列令人信服的实验，证明脲酶是蛋白质。Summer 成功地分离和结晶出脲酶起初得到很多生物化学家的忽略和怀疑。1930 年，John Howard Northrop 从胃蛋白酶商品制剂中分离到了结晶的胃蛋白酶，之后他又结晶了一系列其他的酶，并用更严密的方法证明酶是蛋白质。酶本质的揭示为现代酶学的发展奠定了基础。1946 年，Sumner 和 Northrop 一起荣获诺贝尔化学奖。

推荐网址：

e8-2 本章小结填填看

1. http://en.wikipedia.org/wiki/Enzyme（维基百科有关酶的内容）
2. http://www.lsbu.ac.uk/biology/enzyme/lectures.html（英国伦敦南岸大学提供的完全免费的酶学课程，内容全面，十分有用）

参考文献：

1. Mesecar A D, Koshland D E. A four location model to explain the stereospecificity of proteins. *Nature*, 403: 614–615, 2000.
2. Jeffrey D, et al. Defensive spray of the bombardier beetle: a biological pulse jet. *Science*, 248: 1 219–1 221, 1990.
3. Kofahl R E. The bombardier beetle shoots back. *Creation/Evolution*, 2(3): 12–14, 1981.

第九章　酶动力学

酶动力学（enzyme kinetics）是研究酶促反应速率及其变化规律和影响因素的一门学科。其中涉及的很多名词和基本概念均来自非酶反应动力学。有多种因素可影响酶促反应的速率。在研究某种因素对酶促反应速率的影响时，应该保持其他因素不变，而只改变需要研究的因素。

研究酶促反应动力学具有重要的理论意义和实践意义。这既有助于阐明酶的结构与功能之间的关系，又可为研究酶的作用机理提供有用的数据，除此以外，还有助于寻找最佳的反应条件，了解酶在代谢中的作用以及某些药物作用的机理等。

本章将重点介绍酶的米氏动力学、酶抑制剂作用的动力学、多底物酶的动力学和别构酶的动力学。

第一节　影响酶促反应的因素

一、酶促反应速率和反应类型

Quiz1　在测定一种酶促反应速率的时候，测定[P]的变化一般要比测定[S]的变化更精确，为什么？

酶反应速率与非酶促反应一样，一般都是用单位时间（t）内底物或产物浓度的变化值来表示，常用的单位是 $mol \cdot L^{-1} \cdot s^{-1}$。

对于最简单的单底物和单产物反应 S → P，反应速率用公式可表示为：$v=\frac{d[P]}{dt}=-\frac{d[S]}{dt}=k[S]$，$k$ 为速率常数，v 与[S]成正比，此反应为一级反应；对于同种双分子底物反应 2S → P，$v=\frac{d[P]}{dt}=-\frac{d[S]}{2dt}=k[S]^2$，$v$ 与$[S]^2$ 成正比，反应为二级反应；对于异种双分子底物反应 A + B → P，$v=\frac{d[P]}{dt}=-\frac{d[A]}{dt}=-\frac{d[B]}{dt}=k[A][B]$，反应也是二级反应。

二、影响酶促反应速率的因素

影响酶促反应速率的主要因素包括：酶浓度、底物浓度、反应温度、反应介质的 pH 和离子强度以及有无抑制剂或激活剂的存在等。

（一）酶浓度对酶促反应速率的影响

为了能真实地研究酶浓度对反应速率的影响，底物必须过量存在，以使反应速率不受底物浓度的限制。图 9-1 为测定的结果，从图中可以看出，在一定的 pH 和温度下，当底物浓度大大超过酶浓度时，反应速率与酶的浓度呈正比关系。

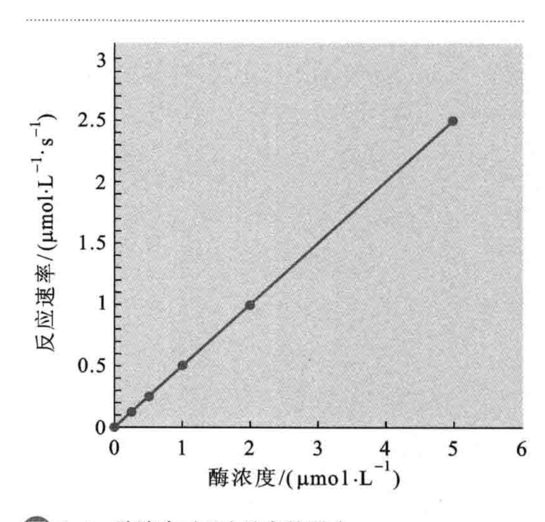
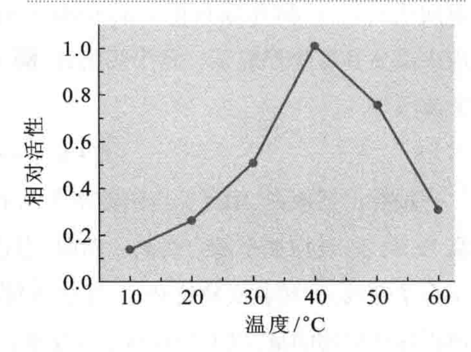


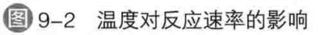
图 9-1　酶浓度对反应速率的影响

（二）温度对酶促反应速率的影响

与大多数化学反应相似，酶促反应速率也会随温度升高而加快，因为温度升高会提高分子之间的碰撞机会。一般而言，温度每升高 10℃，大多数酶的活性增加 50%~100%。然而，酶主要是蛋白质，即使是核酶也会随温度的不断升高而变性。一旦变性，

酶活性会急剧下降。在温度较低时，温度对碰撞机会影响较大，反应速率随温度升高而加快，但温度超过一定数值后，酶受热变性的因素占优，反应速率反而随温度上升而下降，形成倒 V 形曲线（图 9–2）。在此曲线顶点所代表的温度下，酶活性最高，这时的温度被称为酶的最适温度（optimum temperature）。




图 9–2　温度对反应速率的影响

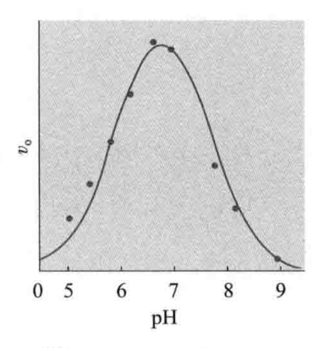


图 9–3　pH 对反应速率的影响

从动物组织提取的酶，其最适温度多在 35~40℃，温度超过 40℃以后，大多数酶开始变性，到了 80℃以上，多数酶的变性已不可逆。然而，从嗜热菌内提取出来的酶对热很稳定，其最适温度较高，如 *Taq* DNA 聚合酶能抵抗 100℃的高温，其最适温度约为 70℃。虽然酶活性随温度的下降而降低，但低温一般不破坏酶。温度回升后，酶又会恢复活性。因此，酶通常储存在 5℃或更低的温度下，只有少数酶会因受冻而失活。临床上低温麻醉就是利用酶的这一性质以减缓组织细胞的代谢速率，提高机体对氧和营养物质缺乏的耐受性，有利于进行手术治疗。

酶的最适温度不是酶的特征性常数，原因是它与反应时间有关，不是一个固定的值。酶可以在短时间内耐受较高的温度；相反，延长反应时间，最适温度便降低。

（三）pH 对酶促反应速率的影响

pH 不仅对酶的稳定性有影响，还对其活性有影响。就后者而言，酶反应介质的 pH 可影响酶分子，特别是酶活性中心上必需基团的解离状况和解离程度，以及催化基团中质子供体或质子受体所需的离子化状态，同时也可影响底物和辅酶（或辅基）的解离状态和解离程度，从而影响到酶与底物的结合。只有在特定的 pH 下，酶、底物和辅酶（或辅基）的解离情况最适于它们相互结合，并发生催化作用，使酶活性最高，这时的 pH 被称为酶的最适 pH（optimum pH）（图 9–3）。最适 pH 和酶的最稳定 pH 不一定相同，和体内环境的 pH 也未必相同。

动物体内多数酶的最适 pH 接近中性，但也有例外，如胃蛋白酶的最适 pH 约 1.8，肝精氨酸酶最适 pH 约为 9.8，溶酶体内的各种水解酶的最适 pH 都偏酸性。

最适 pH 也不是酶的特征性常数，它受底物浓度、缓冲液的种类及浓度以及酶的纯度等因素的影响。溶液的 pH 高于和低于最适 pH 时，都会使酶的活性降低，远离最适 pH 时甚至会导致酶的变性失活。测定酶的活性时，应选用适宜的缓冲液，以保持酶活性的相对恒定。

（四）底物浓度对反应速率的影响

酶动力学研究的核心内容是揭示底物浓度的变化与酶促反应速率之间的关系，而揭示两者之间关系的最好方法也许是用简单的数学方程来描述和概括。为了简化反应系统，需要假定研究的酶无别构效应等特殊的性质，且为单底物和单产物反应。真正的酶虽然很少满足以上假定的条件，但这为理解更为复杂的反应系统提供了便利。

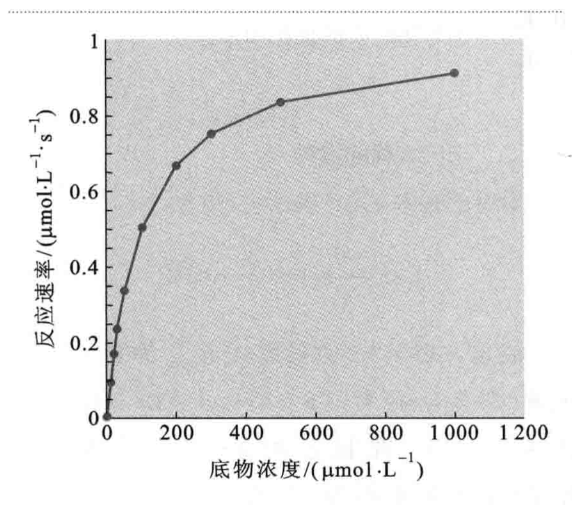


图 9–4　底物浓度对酶反应速率的影响

对于一个正常的非催化的化学反应而言，反应速率与反应物的浓度成正比，以反应速率对反应物浓度作图应该是一条直线。但对于绝大多数酶促反应来说，以反应速率 v 对底物浓度[S]作图，得到的就是双曲线（图 9–4）。

为了解释酶促反应速率与底物浓度之间的

双曲线关系，1888 年瑞典化学家 Savante Arrhenius 提出"酶－底物中间物"假说，他认为，酶在催化反应中需要和底物形成某一种中间物，即酶－底物复合物（enzyme substrate complex，ES），于是反应式需要改写成：

$$E+S \longrightarrow ES \longrightarrow EP \longrightarrow E+P$$

按照上述假说，酶促反应的速率实际上与 ES 的量有关，在一定的酶浓度下，提高底物的浓度会提高 ES 的量，反应速率随之提高。然而，当底物浓度提高到一定水平以后，反应系统中所有的酶分子都结合了底物，即酶被底物饱和了，在这种情况下，ES 达到最大值，反应速率也就达到了最大值，这时即使再提高底物的量，也不会提高反应速率。

酶促反应中酶与底物形成的中间复合物已被许多实验所证实。但由于酶促反应极快，ES 中间物的存留时间极短，需要使用特殊的手段才能检测到它们的存在。1937 年，耶鲁大学的 Kurt G. Stern 将过氧化氢酶（catalase）与其底物过氧化氢的衍生物混合在一起，发现随着反应的进行，酶的光谱发生漂移（spectral shift）。该实验表明，酶首先与底物通过某种方式形成复合物，在反应结束以后，又恢复到原来的状态。后来有人将弹性蛋白酶的晶体在 70% 甲醇溶液中冷却到 -55℃，然后加入其专一性底物——苄氧羰基丙氨酸对硝基苯酯，浸泡一段时间以后，对酶晶体进行 X 射线衍射分析，结果底物浸泡前后的酶差电子密度图清楚地显示，酶与底物形成了酶的酰化中间物。

第二节　米氏反应动力学

酶反应动力学最简单的模型由 Lenor Michaelis 和 Maude Menten 于 1913 年提出，因此又名为 Michaelis-Menten 模型或 M–M 模型。1925 年，G. E. Briggs 和 James B. S. Haldane 使用稳态近似法（steady-state approximation，SSA）提出一个略为复杂的模型，但为了表达对 Michaelis 和 Menten 的开拓性工作的敬意，仍然称之为 M–M 模型。

为了简化反应系统，通常以最简单的单底物和单产物反应为例，并设立 3 个前提条件来推导动力学方程。

一、米氏方程成立的前提

米氏反应动力学需要满足 3 个条件：①反应速率为初速率，此时反应速率与酶浓度呈正比关系，避免了反应产物以及其他因素的干扰；②酶底物复合物处于稳态（steady-state），即[ES]不发生变化 $\left(\frac{d[ES]}{dt}\approx 0\right)$；③符合质量作用定律。

图 9–5　酶促反应过程中的各种变化曲线

二、米氏方程的推导

对于单底物－单产物反应（图 9–5）：

$$E+S \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} ES \xrightarrow{k_2} E+P$$

假定 v_f 表示 ES 的形成速率，v_d 为 ES 的解离速率，那么 $v_f=k_1[E][S]$，而 $v_d=k_{-1}[ES]+k_2[ES]=(k_{-1}+k_2)[ES]$。在稳态时，ES 的形成速率与 ES 的解离速率相等，因此 $v_d=v_f$，即 $k_1[E][S]=(k_{-1}+k_2)[ES]$（①式）。

假定[E_t]表示酶的总浓度，[E]表示游离的酶浓度，[ES]为与底物结合的酶浓度，则[E_t]=[E]+[ES]。于是，①式可变为：

$$k_1([E_t]-[ES])\cdot[S]=(k_{-1}+k_2)[ES]\Rightarrow k_1[E_t][S]-k_1[ES][S]=(k_{-1}+k_2)[ES]$$

$$\Rightarrow k_1[E_t]=k_1[ES][S]+(k_{-1}+k_2)[ES]\Rightarrow k_1[E_t][S]=(k_1[S]+(k_{-1}+k_2))[ES]$$

$$\Rightarrow[ES]=\frac{k_1[E_t][S]}{k_1[S]+(k_{-1}+k_2)}=\frac{[E_t][S]}{[S]+\frac{k_{-1}+k_2}{k_1}}\xrightarrow{k_m=\frac{k_{-1}+k_2}{k_1}}[ES]=\frac{[E_t][S]}{[S]+K_m}$$

由于酶促反应的初速率即是产物形成的速率，$v=k_2[ES]$，于是，

$$v=\frac{k_2[E_t][S]}{[S]+K_m}$$

当[S]→∞，酶被底物饱和，这时的反应速率为最大反应速率 V_{max}，即

$$v=\frac{k_2[E_t]}{1+\frac{K_m}{[S]}}\xrightarrow{[S]\to\infty}v=k_2[E_t]=V_{max}$$

如果将 $k_2[E_t]$换成 V_{max}，则米氏方程可重写成：$v=\frac{V_{max}[S]}{[S]+K_m}$

三、米氏方程的解读和延伸

1. 解读 K_m

K_m 就是米氏常数，为米氏酶的特征常数。人们常常将 K_m 与酶对底物的亲和力相联系，但有时并不确切。事实上，更为确切的说法应该是：对于遵守米氏动力学的反应，K_m 是酶发生有效催化时，对所需底物浓度的一种尺度，即具有高 K_m 的酶比具有低 K_m 的酶需要更高的底物浓度才能到达给定的反应速率。

除了知道 K_m 的字面含义之外，还可以从以下几个方面理解它：

(1) $K_m=\frac{k_{-1}+k_2}{k_1}$，而酶与底物的解离常数 $K_d=\frac{[E][S]}{[ES]}=\frac{k_{-1}}{k_1}$，当 $k_2\ll k_{-1}$ 时，$K_m=K_d$，因此在一定条件下，可以使用它来表示酶与底物的亲和力。一个酶的 K_m 越大，意味着该酶与底物的亲和力越低；反之，K_m 越小，该酶与底物的亲和力越高。

有些酶能催化几种不同的底物发生类似的反应，但与每一种底物的 K_m 都不同。显然，如果将它的各种底物等比例地混合以后再加入酶，那 K_m 最小的底物就最容易与酶结合。

(2) 如果[S]=K_m，米氏方程可转变成 $v=\frac{V_{max}[S]}{[S]+K_m}=\frac{V_{max}[S]}{[S]+[S]}=\frac{V_{max}}{2}$，因此 K_m 是酶反应初速率为 V_{max} 的一半时底物的浓度，它的单位就是浓度的单位。

(3) K_m 可以帮助判断体内一个可逆反应进行的方向。如果酶对底物的 K_m 值小于对产物的 K_m 值，则反应有利于正反应。否则，有利于逆反应。

2. 解读 V_{max}

在特定的酶浓度下，V_{max} 也是一种酶的特征常数，然而，在现实的条件下，一个酶促反应很难达到或者根本就达不到此值。随着底物浓度的增加，v 只能接近此值。如果一个酶促反应的酶浓度发生变化，V_{max} 会随之发生改变。因此严格地说，一个酶促反应的 V_{max} 只有在酶浓度固定在一个值的时候，才是一个常数。

对于大多数酶来说，反应速率随着底物浓度的升高而加快。实际上不管添加多少底物，酶反应速

率从来不会停止增长，只是增长的幅度（图 9–5 中曲线的斜率）越来越小。理论上只有当底物浓度达到无穷大的时候，反应速率才会达到最大值。这就意味着 V_{max} 从来不能被直接测定到，只能通过估算得到。在某些情况下，能得到的最大反应速率实际上远远低于真实的 V_{max} 值，这可能是因为底物的溶解性不好，难以提供很高的底物浓度，也可能是某些酶的活性会被高浓度的底物抑制（参看酶活性的抑制）。

3. 解读 k_{cat}

米氏方程可重写成：$v=\dfrac{k_{cat}[E_t]\cdot[S]}{K_m+[S]}$，这里 $V_{max}=k_{cat}[E_t]$，k_{cat} 整合了在 ES 和 E+P 之间所有反应的速率常数。对于两步反应而言，$k_{cat}=k_2$；对于更复杂的反应，k_{cat} 取决于限速步骤。

k_{cat} 给出了酶被底物饱和以后其催化产物的生成情况。k_{cat} 的单位是时间单位的倒数（例如 s^{-1}），其倒数被认为是一个酶分子"转换"（turn over）一个底物分子所需要的时间。有时 k_{cat} 称为酶的周转数（turnover number）或酶的催化常数，具体是指在单位时间内，一个酶分子将底物转变成产物的分子总数。如果一个酶遵守米氏方程，则 $k_{cat}=k_2=\dfrac{V_{max}}{[E_t]}$。

► 表 9–1　几种酶的动力学参数

酶	底物	$K_m/(mol\cdot L^{-1})$	$k_{cat}/(s^{-1})$	$k_{cat}/K_m/(mol^{-1}\cdot L\cdot s^{-1})$
乙酰胆碱酯酶	乙酰胆碱	9.5×10^{-5}	1.4×10^{4}	1.5×10^{8}
碳酸酐酶	CO_2	1.2×10^{-2}	1.0×10^{6}	8.3×10^{7}
	HCO_3^-	2.6×10^{-2}	4.0×10^{5}	1.5×10^{7}
磷酸丙糖异构酶	3– 磷酸甘油醛	1.8×10^{-5}	4.3×10^{3}	2.4×10^{8}
过氧化氢酶	H_2O_2	1.1	4.0×10^{7}	4.0×10^{7}
胰凝乳蛋白酶	*N*– 乙酰甘氨酸乙酯	4.4×10^{-1}	5.1×10^{-2}	1.2×10^{-1}
	N– 乙酰缬氨酸乙酯	8.8×10^{-2}	1.7×10^{-1}	1.9
	N– 乙酰酪氨酸乙酯	6.6×10^{-4}	1.9×10^{2}	2.9×10^{5}
顺乌头酸酶	顺乌头酸	5.0×10^{-6}	8.0×10^{2}	1.6×10^{8}
	苹果酸	2.5×10^{-5}	9.0×10^{2}	3.6×10^{7}
脲酶	尿素	2.5×10^{-2}	1.0×10^{4}	4.0×10^{5}

在一个细胞里，有成百上千种不同的酶。酶不同，催化效率也会不同。那么，如何衡量和比较不同的酶的催化效率呢？K_m 只能反映一种酶与其底物之间的亲和力。一种酶与底物的亲和力越高，并不能说明这种酶催化的效率就高。事实上，如果酶与底物的亲和力太高，反而不利于催化，因此 K_m 是不能用来显示一种酶的催化效率的。那么，能否用 V_{max} 呢？同样不能，一是因为 V_{max} 在细胞里很难达到，二是因为 V_{max} 随着酶浓度的变化而变化。于是，k_{cat} 就成了衡量一种酶催化效率的非常重要的指标。因为一个酶的周转数越高，就意味着在单位时间内它能将更多的底物转变成产物，它的催化效率也就越高。表 9–1 显示了几种重要的酶的周转数，过氧化氢酶最高，每秒钟 1 个酶分子竟然能将四千万个 H_2O_2 分子分解掉！该酶如此高的催化效率对机体来说是非常重要的，因为一个有氧生物在正常代谢中，会不断产生对机体有毒的 H_2O_2，细胞需要这种高效的"清道夫"来对 H_2O_2 进行解毒。但机体并不需要所有的酶都具有非常高的催化效率，有时候酶的催化效率太高反而是有害的。

Quiz2 如果乙酰胆碱酯酶的 k_{cat} 达到过氧化氢酶的水平，其后果是什么？

4. 解读 $\dfrac{k_{cat}}{K_m}$

$\dfrac{k_{cat}}{K_m}$结合了 k_{cat} 和 K_m，大的 k_{cat} 和 / 或小的 K_m 将给出大的$\dfrac{k_{cat}}{K_m}$值。当$[S]\ll K_m$，即底物浓度很稀的

框 9-1　理论联系实际——甲醇或二甘醇中毒的解毒

全球每年都有人因喝了工业酒精勾兑的假酒中毒致盲或致死的报道。另外，二甘醇作为一种保湿剂，曾广泛用在牙膏当中，只是到了 2007 年 7 月，在中国制造的牙膏先后在巴拿马和美国被指二甘醇含量过高而遭到禁令以后，人们才对它有所了解。但事实上，甲醇和二甘醇本身基本上都没有毒，只是在它们进入人体以后，在人体内乙醇脱氢酶的催化下，被氧化以后才有毒。如甲醇在视网膜中变成甲醛，就必然会影响到视觉的正常功能。

乙醇脱氢酶主要存在于肝细胞，其生理功能显然不是为贪杯者准备的，而是为了帮助机体代谢从食物中带入体内的乙醇，或者由肠道细菌发酵产生的乙醇。在视网膜细胞中，也有乙醇脱氢酶，但它的生理功能是促进视黄醇变为视觉产生必需的视黄醛。

乙醇脱氢酶的底物不止乙醇，还有视黄醇、甲醇、异丙醇和二甘醇，但对乙醇的 K_m 值最小，因此，在同等的条件下，乙醇脱氢酶优先与乙醇结合，并将乙醇脱氢，产生乙醛。如果其他底物与乙醇脱氢酶结合，也可以脱氢产生相应的氧化产物。不过甲醇脱氢产生的甲醛和二甘醇脱氢产生的 2- 羟基乙氧基乙醛对机体是有剧毒的。如果不小心摄入到体内，只要它们不和肝细胞内的乙醇脱氢酶结合，就是安全的，因此当有人不小心摄入少量的甲醇和二甘醇以后，为了防止其中毒，可以让其饮入少量的优质白酒，这实际上就是利用乙醇与乙醇脱氢酶的亲和力较高，而阻止甲醇或二甘醇与乙醇脱氢酶结合转变成有毒的产物，从而使它们有足够的时间从肾排出体外。因此，在某种意义上，酒量越大的人越容易发生甲醇或二甘醇中毒！

时候，米氏方程可转变为 $v=\frac{k_{cat}}{K_m}[E][S]$，这时$\frac{k_{cat}}{K_m}$表现为底物和自由酶之间反应的二级速率常数。该常数十分重要，因为它显示了在有足量酶的存在下，酶和底物能做什么，还允许直接比较酶对不同底物的催化效率，例如表 9-2 中的胰凝乳蛋白酶对三种不同的底物的$\frac{k_{cat}}{K_m}$值差之甚远，这反映了此酶对它三种不同的底物的催化效率是不一样的。因此，$\frac{k_{cat}}{K_m}$也可以用来衡量酶的催化效率，不过更有用的是它还可以用来显示一种酶的完美程度。

作为一个二级速率常数，$\frac{k_{cat}}{K_m}$有一个最可能的极大值，此值由酶和底物的碰撞频率决定，能达到这种“催化境界”的酶可以说已进化得近乎完美！这样的酶促反应速率的限速因素已不是酶本身了，而仅受物理学上无法克服的扩散定律限制了，因为每一次碰撞都会导致反应。如果每一次碰撞导致酶 - 底物复合物的形成，扩散理论预测$\frac{k_{cat}}{K_m}$ 值到达 10^8~10^9 $mol^{-1}\cdot L^{-1}\cdot s^{-1}$。碳酸酐酶、顺乌头酸酶和磷酸丙糖异构酶实际上都接近这个值。但磷酸丙糖异构酶作为糖酵解途径中的一个必需的酶，存在于所有活细胞内，故此酶被公认为当今所有的生命系统中最完美的酶。

总之，$\frac{k_{cat}}{K_m}$对于酶来说是一个非常重要的常数，具体表现在：①它是一个表观二级速率常数；②它可以表示一个酶的催化效率或者完美程度；③它可以反映在较低的底物浓度下，一个酶的催化“表现”；④ $\frac{k_{cat}}{K_m}$的上限 k_2 是由酶与底物扩散到一起的限制因素所决定的。

四、米氏方程的双重性

在底物浓度很低的情况下，即$[S]\ll K_m$，米氏方程可转变为：$v=\frac{V_{max}[S]}{[S]+K_m}=\frac{V_{max}[S]}{K_m}$。这时反应

速率与底物浓度成正比，符合一级动力学；在底物浓度很高的情况下，即[S]>>K_m，米氏方程可转变为：$v=\frac{V_{max}[S]}{[S]+K_m}=\frac{V_{max}[S]}{[S]}=V_{max}$。这时酶反应速率接近最大值，如果继续增加底物的浓度，v 不可能继续增加，反应速率与底物浓度的关系符合零级动力学。

五、米氏方程的线性转换

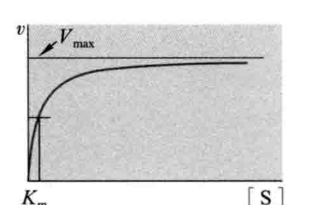


图 9-6 米氏酶反应速率与底物浓度的关系曲线

直接使用米氏方程中的 v 对[S]作图得到的是一条双曲线，从图中也能得到 K_m 值（反应速率为最大值一半时底物的浓度）和 V_{max} 值（图 9-6），然而，由于实验误差的客观存在，只要出现任何可见的误差使数据点偏离真实的位置，就会很难画出一个很完美的曲线，而在同样的条件下，画直线更容易。因此，将米氏方程进行线性化处理就显得十分必要。

1. Lineweaver-Burk 作图

Lineweaver-Burk 作图可谓是最普及、最常见的确定酶动力学常数的方法，该作图法需要对米氏方程作以下转换：

$$v=\frac{V_{max}[S]}{[S]+K_m} \Rightarrow \frac{1}{v}=\frac{[S]+K_m}{V_{max}[S]}=\frac{1}{V_{max}}+\frac{K_m}{V_{max}}\cdot\frac{1}{[S]}$$

转换以后，以$\frac{1}{v}$对$\frac{1}{[S]}$作图，将得到一条直线（图 9-7）。由于这种作图法的变量变成了 v 和[S]的倒数，因此又称为双倒数作图（double reciprocal plot）。

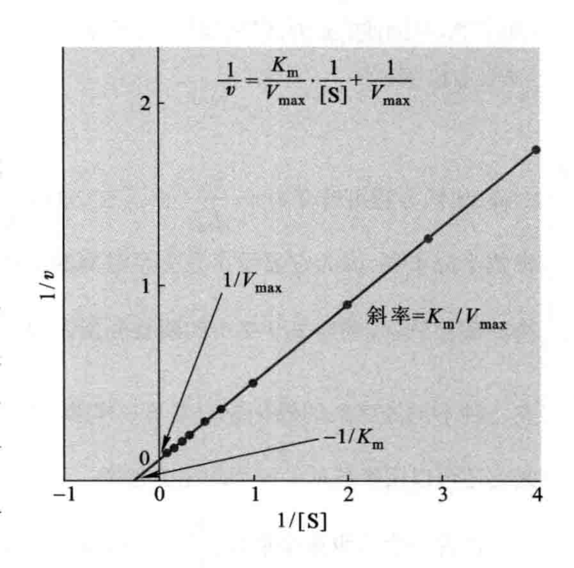


图 9-7 米氏酶的双倒数作图

使用此作图法，不需要直接测量 V_{max}，从图中就能得到 V_{max} 和 K_m，但它也许是最不准确的计算酶动力学常数的方法。

如果实验数据没有误差，双倒数作图无疑就是最佳的确定酶动力学常数的方法。然而，实验的误差总是存在的，而双倒数作图不能很好地处理这些误差，特别是在底物浓度较低的情况下的误差会导致$\frac{1}{v}$出现更大的误差，这时作出来的点分散得十分严重。

为了更好地说明双倒数法的缺陷，让我们假定一个遵守米氏动力学方程的酶，同时得到了几组含有偶然误差的数据，如果使用微软的 Excel，v 对[S]作图的结果如图 9-8(1)，而使用同样的数据用双倒数作图的结果如图 9-8(2)，图中的实心圆点代表无误差的数据。

比较两图不难发现，同样具有误差的数据，特别是在低底物浓度下得到的数据（图中的右边）在双倒数法中得到的点分散严重，换句话说，双倒数法扩大了误差。

e9-1 其他线性化作图法

Quiz3 哪一种作图法的误差最小？为什么？

除了经典的双倒数作图以外，还可以对米氏方程进行其他形式的变换，从而得到另外两种线性化作图方法—— Eadie-Hofstee 作图和 Hanes-Wolff 作图。此外，还有一种直接线性化作图。这 3 种作图法各有利弊。

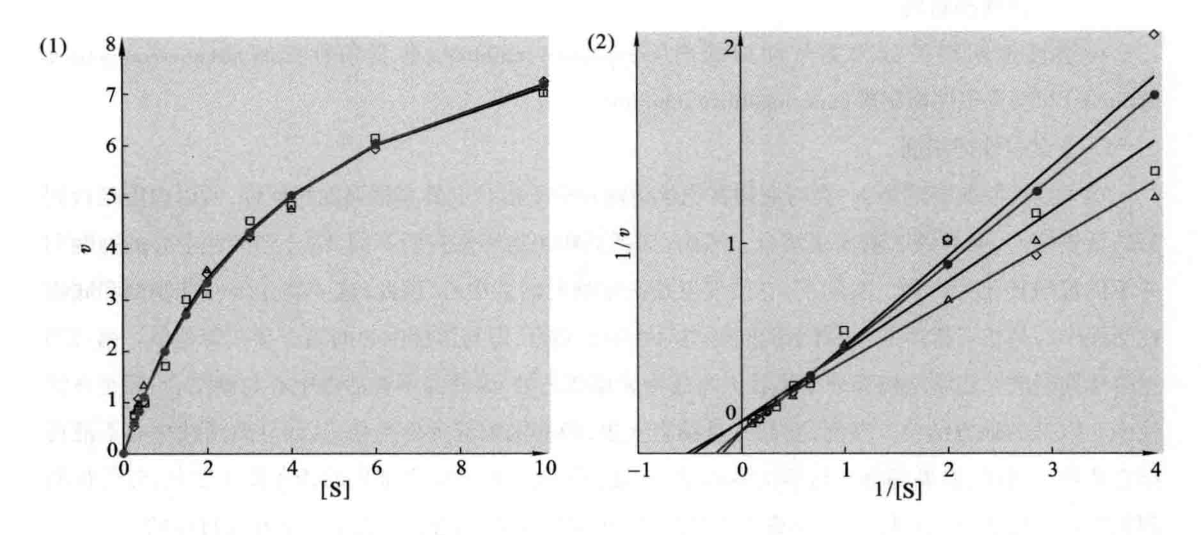


图 9-8　双倒数作图对 v 和[S]误差的扩大

第三节　酶抑制剂作用的动力学

有两种情况可导致酶活性的降低或丧失：一种是由酶分子变性引起的，另外一种是由抑制剂的作用造成的。酶动力学分析的一个重要用途是研究酶抑制剂的作用。酶抑制剂泛指那些通过与酶分子的结合而降低反应速率、抑制酶活性的物质。它们可能是在细胞中自然存在，用来控制代谢反应速率的调节物，也可能是人工合成的有机的或无机的试剂。许多毒性化合物实为酶的抑制剂，在细胞内能抑制关键的代谢反应，这些毒性抑制剂对某种或某类有机体有特异性，因此可用做抗生素、杀虫剂和除草剂等。许多抗病毒和抗癌药物也是特定的酶抑制剂（表 9-2）。

表 9-2　几种常见的酶抑制剂药物

药物名称（中文）	药物名称（英文）	靶酶	医用或药用
阿司匹林	aspirin	前列腺素合成中的环加氧酶	消炎
青霉素	penicillin	肽聚糖转肽酶	抗生素
甲氨蝶呤	methotrexate	二氢叶酸还原酶	抗肿瘤
叠氮胸苷	azidothymine（AZT）	HIV 逆转录酶	艾滋病治疗
利托那韦	ritonavir	HIV 蛋白酶	艾滋病治疗
万艾可	viagra	cGMP 磷酸二酯酶 -5	勃起功能障碍（ED）

抑制剂能够以不同的方式作用于酶，而酶动力学是区分各种作用方式的主要手段。

根据抑制方式，酶的抑制剂可分为可逆性抑制剂（reversible inhibitor）和不可逆性抑制剂（irreversible inhibitor）。前者以次级键与酶可逆结合，这些键形成得快，断裂得也快，结果是抑制作用的效果来得快，但并不能使酶永久性失活，使用透析或超滤就可去除它们，让酶恢复活性；后者也称为酶灭活剂（inactivator），它们以强的化学键（通常是共价键）与酶不可逆结合，结合一个，灭活一个，导致酶的有效浓度降低，酶一旦失活就不可逆转。如果想恢复酶的活性，唯一的手段只能补充新酶。与可逆性抑制剂不同的是，不可逆性抑制剂需要更长的时间与酶起反应，因为共价键形成较慢。因此，不可逆性抑制剂通常表现出对时间的依赖性，即抑制的效果随着与酶接触时间的延长而增强。

一、可逆性抑制剂

可逆性抑制剂又分为竞争性抑制剂(competitive inhibitor)、非竞争性抑制剂(non-competitive inhibitor)和反竞争性抑制剂(uncompetitive inhibitor)。

(一) 竞争性抑制剂

竞争性抑制剂有两类:一类与底物在化学结构和分子形状上具有很强的相似性,所以也能结合到酶的活性中心,而导致底物无法结合。然而,抑制剂和底物毕竟有所不同,结合到活性中心的抑制剂并不能被酶转化为产物。实际上,它们仅仅是占据酶的活性中心,因此,这一类竞争性抑制剂和底物在活性中心是相互排斥的,两者不能同时与活性中心结合,即与活性中心的结合存在着竞争。第二类竞争性抑制剂在化学结构和分子形状上与底物无相似之处,因此并不在活性中心与酶结合,而是在活性中心以外的地方结合。然而,它们一旦结合上去,酶的构象就发生变化,从而导致活性中心不能再结合底物。同样,如果底物先与活性中心结合,就会导致抑制剂结合部位的构象发生变化,致使抑制剂无法再与酶结合。因此,这一类竞争性抑制剂与底物在和酶结合这一点上也是相互排斥的。

图 9-9 细菌的二氢叶酸合成酶的正常底物和竞争性抑制剂的化学结构

作为竞争性抑制剂两个最典型的例子,一个是磺胺(sulfanilamide)类药物(图 9-9),另一个是丙二酸(图 9-10)。前者是对氨基苯甲酸(para-aminobenzoic acid,PABA)的类似物,能竞争性抑制细菌内以 PABP 为底物的一系列酶(如二氢叶酸合成酶)的活性,从而阻止 PABA 作为原料合成细菌所需要的叶酸,进而抑制细菌核苷酸和核酸的合成,起到抗菌作用;后者是琥珀酸的类似物,能竞争性抑制琥珀酸脱氢酶的活性,从而阻断三羧酸循环。

图 9-10 琥珀酸脱氢酶的正常底物和竞争性抑制剂

既然任何一种抑制剂都能够减慢反应速率,那么,在它们作用于酶以后,必然会改变酶的动力学行为。不同类型的抑制剂影响的效果不一定相同,借此可以区分抑制剂的类型。

研究抑制剂作用动力学的一种途径是,测定在有和无抑制剂的条件下的酶促反应速率。如图 9-11 所示,就竞争性抑制剂而言,在底物浓度较低的时候,抑制剂能显著地降低反应速率,但如果底物浓度非常高,速率改变得就不明显。

图 9-11 竞争性抑制剂对酶促反应速率的影响

图 9-12 竞争性抑制剂与酶结合的反应式

在有竞争性抑制剂存在的情况下,酶既能与抑制剂结合形成酶抑制剂复合物(EI),也能与底物结合形成酶与底物复合物(ES)(图 9-12)。这两种结合反应都是非常快的,而且也是可逆的,它们之间存在着平衡。毫无疑问,平衡的位置取决于底物与抑制剂之间的浓度差:在抑制剂浓度非常高的情况下,几乎所有的酶分子都与抑制剂形成 EI,无游离的酶分子与底物形成 ES,这时酶活性差不多完全被抑

制；相反，在底物浓度非常高的条件下，几乎所有的酶分子与底物形成 ES，无游离的酶分子与抑制剂形成 EI，这时抑制剂在与底物的竞争中"一败涂地"。

采用稳态近似法，很容易推导出在有竞争性抑制剂情况下的米氏方程：

$$v=\frac{V_{max}[S]}{\alpha K_m+[S]},\quad K_I=\frac{[E][I]}{[EI]},\quad \alpha=\left(1+\frac{[I]}{K_I}\right)。$$

方程中的 K_I 为酶与抑制剂的解离常数，$\alpha=\left\{1+\frac{[I]}{K_I}\right\}$。当没有抑制剂的时候，$\alpha$ 为 1。

根据上述修改后的米氏方程，容易得出：在有抑制剂的时候，酶与底物的表观 K_m 值提高 α 倍，其大小与抑制剂的浓度有关，但反应的 V_{max} 不变。显然，在存在抑制剂的情况下，某些酶分子以自由的形式存在，某些与抑制剂形成 EI。前者与底物以正常的亲和力结合，后者与底物的亲和力为零，即完全不能与底物结合。K_m 反映的是反应混合物内酶与底物的总的亲和力，因此，在有抑制剂的时候，酶与底物的亲和力应该下降，即表观 K_m 值提高。

V_{max} 是在非常高的底物浓度下的反应速率。在非常高的底物浓度下，抑制剂根本不是底物的"竞争对手"，这时候的反应速率同样可以达到原来的最大值。

在不同的抑制剂浓度下，进行双倒数作图，可以清楚地显示出竞争性抑制剂不能改变 V_{max}，但能降低酶与底物的总的亲和力，即提高表观 K_m。抑制剂浓度越高，表观 K_m 提高得越大[见图 9-11(2)]。

（二）非竞争性抑制剂

非竞争性抑制剂在活性中心以外与酶结合，并改变活性中心的构象，这一点似乎和与底物结构无相似性的竞争性抑制剂很像，但是，非竞争性抑制剂并不阻止底物在活性中心与酶结合，只阻止底物转变成产物。

既然非竞争性抑制剂和底物能够同时与酶结合形成酶-底物-抑制剂三元复合物(EIS)(图 9-13)，这就意味着抑制剂在高浓度或低浓度底物下能够等效地发挥抑制作用。虽然有两条形成酶-抑制剂-底物三元复合物 EIS 的途径，但是最后的结果是一样的，都是形成无活性的复合物。采用稳态近似法，也容易推导出在有非竞争性抑制剂情况下的米氏方程：

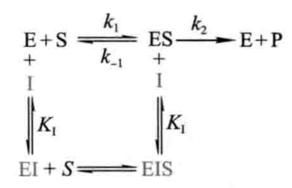


图 9-13 非竞争性抑制剂与酶结合的反应式

$$v=\frac{[S]}{[S]+K_m}\times\frac{V_{max}}{\left(1+\frac{[I]}{K_I}\right)}$$

曲线作图和双倒数作图的结果参看图 9-14。作图结果显示，典型的非竞争性抑制剂不影响酶与底物的亲和力，也就不改变酶的 K_m。但这样的非竞争性抑制剂较为少见，更多的是会降低酶与底物的亲和力，从而导致 K_m 升高。鉴于后者同时具有竞争性和非竞争性的部分性质，这样的非竞争性抑制剂又称为混合型抑制剂(mixed inhibitor)。

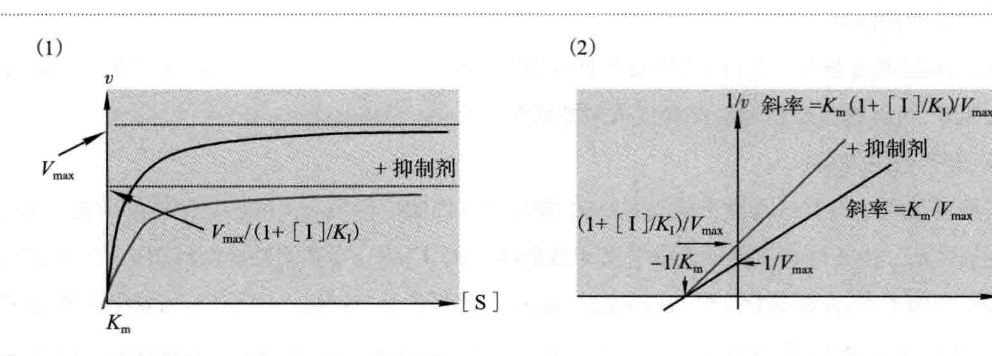


图 9-14 非竞争性抑制的正常作图和双倒数作图

不论是典型的非竞争性抑制剂，还是混合型抑制剂，它们在任何底物浓度下都有抑制作用，因此都会降低 V_{max}。

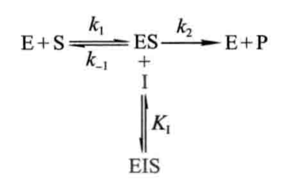


图 9-15 反竞争性抑制剂与酶结合的反应式

（三）反竞争性抑制剂

这是一类只能与酶和底物复合物（ES）结合，但不能与游离的酶结合的抑制剂。一旦与 ES 结合，就导致与活性中心结合的底物不再能够转变为产物（图 9-15）。反竞争性抑制剂之所以只能与 ES 结合，可能是因为底物本身直接参与抑制剂的结合，也可能是因为与底物结合将导致原来不能结合抑制剂的位点构象发生改变，转变成能够结合抑制剂的构象。抑制剂抑制的机理可能是因为它的直接作用，也可能是因为其结合导致活性中心的构象发生了变化。

反竞争性抑制剂的存在可能只有理论上的可能，因为迄今为止，还没有文献报道过现实中有哪一种酶受到一种反竞争性抑制剂的作用，因此，有关它们的描述仅有理论上的价值。由于反竞争性抑制剂只能与 ES 结合，因此在底物浓度很低的时候，酶几乎都处于游离的状态，其抑制作用可以说是微乎其微。而如果底物浓度很高，大多数酶就处于 ES 状态，这时候抑制剂是有效的。

同样，使用稳态近似法可推导出在有反竞争性抑制剂时的米氏方程：

$$v=\frac{[\mathrm{S}]}{[\mathrm{S}]+K_m\Big/\left(1+\frac{[\mathrm{I}]}{K_I}\right)}\times\frac{V_{max}}{\left(1+\frac{[\mathrm{I}]}{K_I}\right)}$$

其曲线作图和双倒数作图结果如图 9-16，图中的结果显示，反竞争性抑制剂的存在能降低酶的表观 K_m。K_m 的降低有点出乎意料，对此的解释是：抑制剂能有效地与 ES 结合，形成 EIS 三元复合物，从而减少了 ES 浓度。根据化学平衡理论，这相当于把酶与底物结合反应的平衡拉向右边，即有利于酶与底物的结合，从而导致酶与底物结合的亲和力的增加，即表观 K_m 降低。

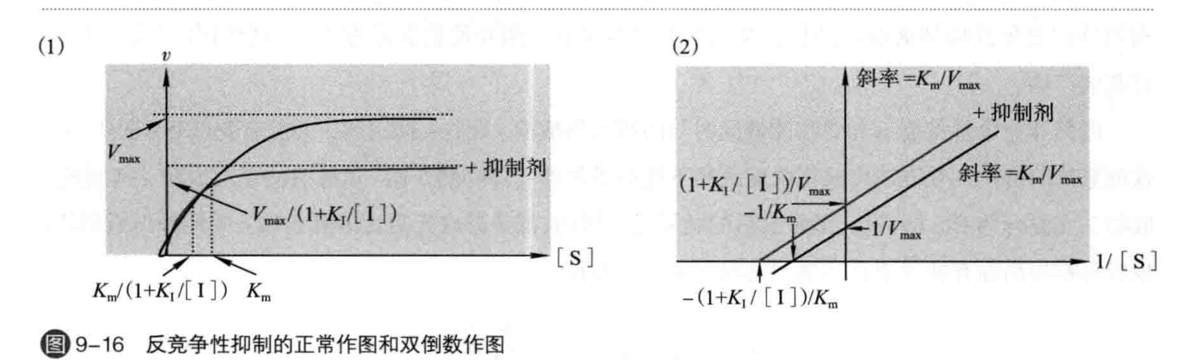


图 9-16 反竞争性抑制的正常作图和双倒数作图

大量底物的存在并不能“压倒”抑制剂，反而有利于抑制剂与 ES 的结合，从而抑制酶促反应，因此反竞争性抑制剂会降低 V_{max}。以上介绍的抑制剂都是来自于一个酶促反应体系之外。事实上，有时一个酶促反应体系内的物质也可能对酶的活性产生抑制，如产物抑制（product inhibition）和过量底物抑制（excess substrate inhibition）。

e9-2 产物抑制与过量底物抑制

二、不可逆性抑制剂

不可逆性抑制剂可分为 4 类，即基团特异性抑制剂（group specific reagent）、底物类似物（substrate analogue）抑制剂、过渡态类似物抑制剂和自杀型抑制剂（suicide inhibitor）。

（一）基团特异性抑制剂

Quiz4 不可逆抑制剂抑制的动力学性质与哪一种可逆性抑制剂相似？哪一种不可逆性抑制剂作用的特异性最低？

这一类抑制剂在结构上与底物无相似之处，但能共价修饰酶活性中心上的必需侧链基团而导致酶活性不可逆的失活。由于许多氨基酸残基含有亲核侧链基团，因而充当基团特异性抑制剂的一般是亲电试剂。常见的例子有：有机磷化合物（organophosphorou）、碘代乙酸（iodoacetamide）和环氧化物等，其中有机磷化合物包括甲基氟磷酸异丙酯（sarin，沙林）和二异丙基氟磷酸（diisopropylfluorophosphate，DIPF）。

DIPF 是最典型的一例，它能够修饰多种酶活性中心上的 Ser 残基的羟基，例如胰凝乳蛋白酶和乙酰胆碱酯酶（acetylcholinesteras），从而导致这些酶活性的丧失。胰凝乳蛋白酶的 Ser195 残基受到 DIPF 的修饰后并失活（图 9-17），乙酰胆碱酯酶活性中心上的 Ser 受到沙林毒气的修饰后同样失活。

图 9-17　DIPF 对胰凝乳蛋白酶活性的抑制

由于乙酰胆碱酯酶在神经递质乙酰胆碱(acetylcholine)的代谢中(水解其成为胆碱和乙酸)起重要作用,因此它的抑制将导致乙酰胆碱的积累,以致肌肉过分收缩,导致肌肉痉挛。死亡可能因为喉痉挛而随时发生。1995 年 3 月 20 日,日本“奥姆真理教”(Aum Shinrikyo)的一名恐怖分子在东京的 5 个地铁站里放置沙林,结果造成 12 人死亡以及 5 千多人受伤。此次沙林毒气案,是世界上首次出现利用化学武器进行的恐怖活动。针对有机磷中毒者的解毒是有可能的。这需要使用亲核性更强的试剂(如解磷定——2- 甲醛肟吡啶碘甲烷盐),使被共价修饰的 Ser 残基的侧链羟基恢复自由。碘代乙酸能修饰多种酶(如 3- 磷酸甘油醛脱氢酶)活性中心上的巯基,从而导致这些酶活性被完全抑制。

(二) 底物类似物抑制剂

这一类抑制剂在结构上分为两个部分:一个部分类似于底物,抑制剂正是通过这个部位“骗过”酶分子,结合到活性中心,锁定抑制的对象;另外一个部分含有反应性基团,在抑制剂进入活性中心以后,可以不可逆地修饰上面的必需基团,导致酶活性的丧失。

这一类抑制剂与竞争性抑制剂的差别在于后者缺乏反应性基团。在酶作用机理的研究中,巧妙地利用这一类抑制剂可以对酶分子的活性中心进行亲和标记(affinity label),以确定反应的必需基团。例如,甲苯磺酰苯丙氨酰氯甲酮(tosyl-L-phenylalanine chloromethyl ketone,TPCK)为胰凝乳蛋白酶的底物类似物,在与胰凝乳蛋白酶的活性中心结合以后,便共价修饰其中的 His 残基,导致酶活性的抑制(图 9-18);再如,甲苯磺酰赖氨酰氯甲酮(tosyl-L-lysine chloromethyl ketone,TLCK)为胰蛋白酶的底物类似物,在与胰蛋白酶的活性中心结合以后,共价修饰活性中心的 His 残基,导致酶活性的丧失。还有,磷酸溴丙酮(bromoacetol phosphate)是磷酸丙糖异构酶的底物磷酸二羟丙酮的类似物,它与磷酸丙糖异构酶的活性中心结合以后,会共价修饰活性中心的 Glu,致使酶活性受到抑制(图 9-19)。

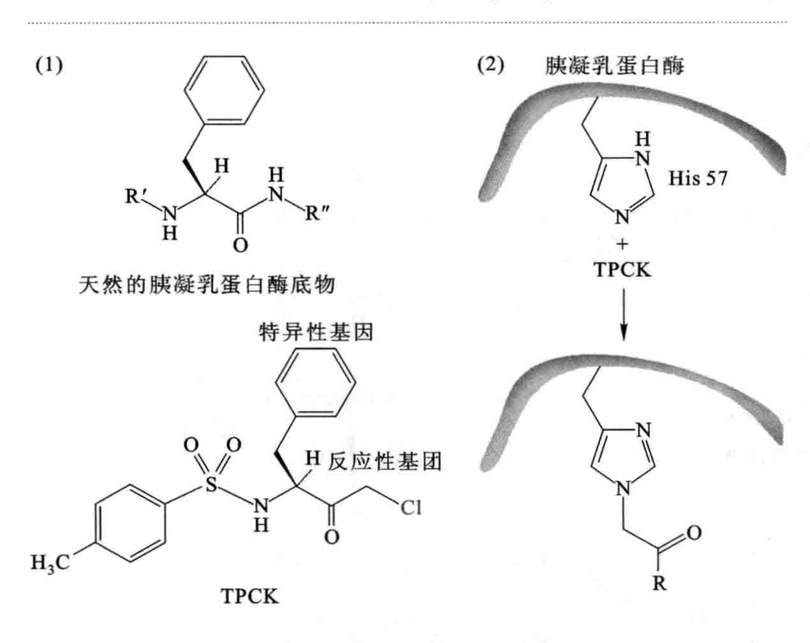


图 9-18　甲苯磺酰苯丙氨酰氯甲酮对胰凝乳蛋白酶活性中心 His 的亲和标记

图 9-19　磷酸溴丙酮对磷酸丙糖异构酶活性中心 Glu 的亲和标记

(三) 过渡态类似物抑制剂

这一类抑制剂与酶促反应的过渡态极为相似,它们

在化学结构和分子形状上与酶的活性中心十分般配，能够以极高的亲和力与活性中心结合，从而导致底物无法进入而使得酶活性受到不可逆性抑制（参看第十章“酶的催化机理”）。有些生物使用天然的过渡态类似物来抑制酶的活性。例如，胰腺细胞里有一种胰蛋白酶的过渡态类似物叫胰胰蛋白酶抑制剂（pancreatic trypsin inhibitor，PTI），它的功能就是抑制在胰腺细胞内提前激活的胰蛋白酶的活性，保护细胞，防止细胞发生自溶（参看第十二章“酶活性的调节”）。

（四）自杀型抑制剂

自杀型抑制剂是受酶本身来激活的不可逆性抑制剂。这类抑制剂与第二类不可逆性抑制剂有点类似，但差别在于其反应性基团是潜在的。只有在与酶结合以后，反应性基团才会受到酶的催化而激活，转而修饰酶的必需基团并导致酶活性的丧失。由于它们依赖于酶正常的催化机理来导致酶的失活，因此被称为机理型抑制剂（mechanism-based inhibitor）；又由于它们“冒充”底物与酶结合并受到酶的激活而抑制酶活性，因此又被称为特洛伊木马（trojan horse）抑制剂；还由于它们是依赖于酶的催化而激活，因此还被称为 k_{cat} 抑制剂。

此类抑制剂具有 3 个重要的特征：①没有酶，无化学活性；②必须受到靶酶的激活；③与酶反应的速率快于它与酶解离的速率。

以 *N*,*N*– 二甲基炔丙胺（*N*,*N*–dimethylpropargylamine，DMPA）为例，它作为单胺氧化酶（monoamineoxidase，MAO）的自杀型抑制剂，在与酶结合以后被氧化，反过来共价修饰酶的黄素辅基（N_5 的烷基化），导致酶活性的不可逆抑制（图 9–20）。由于 MAO 在体内能催化多巴胺（dopamine）和血清素（serotonin）等神经递质的脱氨，促进它们在脑内水平的下降，而帕金森病和抑郁症分别与低水平多巴胺和血清素有关，因此可以使用 DMPA 来提高多巴胺和血清素的水平，从而达到治疗这两种疾病的目的。

Quiz5 为什么过渡态类似物抑制剂和自杀型抑制剂是制药公司作为药物筛选的首选？

图 9–20 *N*,*N*– 二甲基炔丙胺对单胺氧化酶的自杀型抑制

第四节 多底物反应动力学

现实中的单底物/单产物反应实为罕见，更多的是双底物/双产物反应，三底物/三产物反应也很普遍，偶尔还有四底物反应。

底物和产物数目的增加并不意味着应用在单底物/单产物反应中的动力学分析方法就淘汰了。实际上，酶学研究者在研究多底物反应时，照样可以很好地利用在单底物反应中使用的方法，只需要改变其中的一种底物的浓度，而保持其他底物浓度的恒定。但若全然不顾其他底物对反应的影响，就会遗漏掉多底物酶特有的一些有价值的信息。

一、多底物反应中的一些专门术语

多底物动力学比单底物动力学要复杂，需要用更多的术语来进行描述。

首先，底物和产物越多，需要代表它们的符号就越多。通常使用A、B、C和D代表不同的底物，P、Q、R和S代表不同的产物。底物和产物的数目分别用Uni、Bi、Ter和Quad代表1、2、3和4。例如，单底物/三产物反应可表示为“Uni-Ter”。有时可以用反应性(reactancy)来表示底物或产物的多少：一个反应的反应性被定义为沿着一个特定的方向发生反应的底物数目。以一个Bi-Ter反应为例，其正反应的反应性为2，逆反应的反应性为3。其次，在多底物反应中，酶有多种形式，这也需要更多的字母符号来表示。用来表示酶的符号有E、F和G，其中E总是表示游离的酶。至于涉及的抑制剂一般使用I和J表示。

至于酶和底物或酶和产物形成的复合物仍旧使用过渡态复合物(transitory complex)的名称，因为它们存留的时间很短，要么重新解离，要么经历共价变化，从而导致底物转变为产物或产物转变为底物。然而，在多底物反应中，会产生多种类型的过渡态复合物。以Bi-Bi反应为例，有3种可能的酶底物复合物——EA、EB、EAB和3种可能的酶产物复合物——EP、EQ、EPQ。此外，某些酶促反应还可能形成酶-底物-产物的三元复合物，如EAP和EBQ。

在单底物反应系统中，酶的活性中心一旦有一个底物分子结合上去，就能发生反应，根本没有另一个底物分子、产物或抑制剂结合的空间；而在多底物反应中，酶能同时结合多个底物。显然，一个底物分子不能完全填满活性中心。中心复合物(central complex)专指活性中心被底物或产物完全填满的过渡态复合物(通常用括号以示区别)，像Bi-Bi反应中的EAB和EPQ就属于中心复合物，而EA、EB、EP和EQ则不属于。

二、多底物反应的动力学机制

一个多底物酶的动力学机制仅仅是指底物与酶结合的次序和产物从酶释放的次序，它不应该与酶和底物相互作用的化学反应机制混为一谈。一般而言，动力学机制可以分为序列机制(sequential mechanism)和乒乓机制(ping-pong mechanism)。

在研究一个多底物反应的时候，只需将其中一种底物的浓度固定，就完全可以按照一个单底物反应来处理，预期的结果会和单底物反应系统相同，相关的动力学常数同样可以计算出来。如果改变浓度固定的底物的浓度，即让它在不同的固定浓度下，重复实验步骤，测定浓度可变的底物与酶反应速率的关系，可以观察到酶促反应速率会改变。固定底物的浓度越高，反应速率就越高。于是，每一个固定的底物浓度，就可以得到一组动力学数据。使用任意一种作图法，可以得到一系列曲线，由此可计算出与浓度变化的底物有关的动力学常数。而根据反应的动力学行为，区分序列反应和乒乓反应是完全可能的。

e9-3 多底物反应的动力学

第五节　别构酶的动力学

以上所有关于酶动力学的讨论都是基于米氏方程，所涉及的酶促反应速率对底物浓度的作图都呈双曲线。然而，生物体内还有另一类酶偏离米氏动力学，这一类酶属于所谓的别构酶(allosteric enzyme)，它们的一些性质与非酶的别构蛋白(如血红蛋白)相似。

一、别构酶的性质

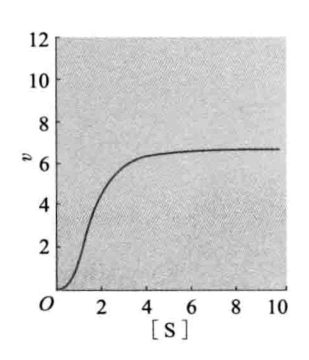


图 9-21　典型的别构酶催化的反应速率与底物浓度的关系曲线

Quiz6　为什么细胞内别构酶的底物浓度通常在对应于S形曲线最陡的点附近?

一种典型的别构酶具有以下性质：

1. 反应速率－底物浓度曲线一般为S形

S形曲线与双曲线最重要的区别是：在底物浓度本来就不高的情况下，提高底物浓度只能引起反应速率幅度极小的增加，这时候曲线的斜率很低；在稍高的底物浓度下(图 9-21 中大于 0.5 单位)，底物浓度的增加会导致反应速率的急剧升高，这时的曲线斜率较高；在底物浓度很高的情况下，曲线平缓，实际上已接近双曲线，反应速率趋于 V_{max}。

S形曲线显示了底物与酶结合的正协同性。在底物浓度很低的时候，只有少数酶活性中心与底物结合，这时底物与酶的亲和性很低，提高底物浓度也只能导致反应速率很小的增加。然而，随着更多的底物与酶结合，正协同效应开始起作用，致使酶与底物的亲和性大增，反应速率随之猛升。当底物浓度提高到一定水平的时候，别构酶就像双曲线酶一样被底物饱和，速率接近 V_{max}。

然而，并不是所有的别构酶的反应速率对底物浓度作图总是S形曲线。像一些多底物酶，也许对某一种底物表现正协同性，但对其他底物无正协同性。显然，这类别构酶的反应速率并非对所有底物的浓度作图都表现为S形曲线。此外，某些别构酶对底物表现的是负协同性，这种情形下的速率对底物浓度的作图也不是S形曲线。

2. 具有别构效应物

别构酶除了含有活性中心以外，还有别构中心。这是别构酶名称的由来，也是判断一种酶是不是别构酶的主要标准。别构中心是指底物以外的分子结合的位点，这些分子被统称为别构效应物(allosteric effector)。其中起激活酶活性的物质称为别构激活剂(allosteric activator)，相反，起抑制作用的称为别构抑制剂。通过别构效应物调节酶活性是细胞代谢调控的重要手段之一，而别构效应物的存在可以改变一个典型的对底物呈正协同性别构酶的动力学行为。

图 9-22 为一个典型的别构酶分别在有无别构效应物的条件下，其反应速率对底物浓度的曲线，中央曲线是在没有效应物存在的情况下得到的，它是一个典型的S形曲线。最上面的两条曲线是在别构激活剂存在的时候得到的，可见激活剂可以提高任何底物浓度下的反应速率。最下面的两条曲线是在抑制剂存在下得到的，可见抑制剂可以降低任何底物浓度下的反应速率。然而，如果仔细比较激活剂和抑制剂对S形曲线走势的影响，就会发现，抑制剂加强曲线的S形，拉长S形曲线的"趾部"，而激活剂具有相反的效果。事实上，在高水平的激活剂存在下，S形曲线会转变成双曲线。由此可以看出，别构抑制剂能增强酶对底物的正协同性，而激活剂则削弱酶对底物的正协同性。

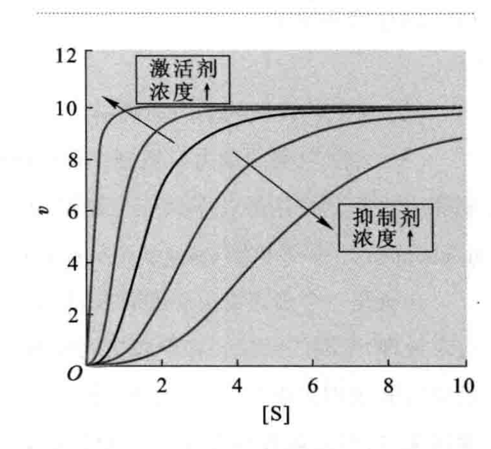


图 9-22　激活剂或抑制剂对别构酶活性的影响

显然，上图的所有曲线都趋向同一个 V_{max}，意味着这一种别构酶的别构效应物通过改变酶与底物的亲和性，即 K_m 值来调节酶活性。这样的系统称为K系统(K-system)。实际上，还有一类别构酶的抑制剂通过改变酶的 V_{max} 而起作用，这样的系统称为V系统(V-system)。

3. 对竞争性抑制的作用表现为双相应答(biphasic response)

除了别构抑制剂以外,别构酶还可能像其他非别构酶一样受到竞争性抑制剂的作用,典型的竞争性抑制剂是通过模拟底物的化学结构起作用的。但对于一个具有正底物协同性的别构酶来说,如果一种竞争性抑制剂在结构上与其底物过分相似,这种抑制剂就可以像底物一样诱发正协同效应的发生。在这样的情况下,低浓度的竞争性抑制剂能够提高酶与底物的结合能力,反而可以提高反应速率(似乎作为激活剂);而高浓度的抑制剂则以通常的方式减慢反应速率。这样的竞争性抑制剂对别构酶活性的双面影响称为双相应答。

4. 温和变性可导致别构效应的丧失

酶的变性是指酶三维结构的破坏。一旦酶变性,其催化活性就会丧失。导致酶变性的因素就是导致蛋白质变性的因素,包括热、极端 pH 和化学变性剂等。受到任何一种变性因素的温和作用,别构酶通常能够保持催化活性,但其别构性质和底物协同性会丧失。这说明三维构象对酶的别构性质同样重要,并且别构性质对变性剂的作用似乎更为敏感。

5. 通常是寡聚酶

别构酶与别构蛋白具有一个共同的性质,就是一般为多亚基蛋白,具有四级结构,亚基之间以次级键相连。到目前为止,发现具有别构效应的单体酶极为少见,丙酮酸 -UDP-*N*- 乙酰葡糖胺转移酶(pyruvate-UDP-*N*-acetylglucosiamine transferase)为一例。

Quiz7 你认为单体别构酶对底物的结合会有协同性吗?为什么?

寡聚别构酶又分为同源寡聚酶和异源寡聚酶,前者由相同的亚基组成,每一个亚基既有活性中心,又有别构中心,后者由不同的亚基组成,活性中心和别构中心分属不同的亚基。

多亚基结构对底物结合的协同性的产生至关重要。一个别构酶会含有几个活性中心,其最简单的布局是每一个亚基含有一个活性中心,每一个活性中心都能行使相同的催化功能。各活性中心的相互作用是底物协同性产生的原因。对于一个典型的具有正底物协同性的别构酶来说,一个底物分子与其中的一个活性中心的结合会诱发其他活性中心的构象发生变化,致使其他活性中心更容易与底物结合;而一个具有负底物协同性的别构酶正好相反,它的一个活性中心与底物结合以后,会降低其他活性中心与底物的亲和力。

亚基之间由于通过次级键结合,因此既容易解离,又容易重新聚合。于是,一个别构酶在溶液中有完整的寡聚体和单体两种形式,且两者处于动态平衡之中。在多数情况下,单体可能无催化活性。而一些复杂的别构酶在完整的寡聚体和单体之间还存在某些中间物,各种配体与酶的结合通常会改变上述平衡。

6. 与非别构酶相比,别构酶占少数。

关于别构机理请参看酶活性调节。

二、S 形曲线和 Hill 方程

既然米氏方程给出的是双曲线,它就不再适用于具有底物协同性的呈 S 形曲线的别构酶,但有一个与米氏方程关系密切的 Hill 方程却能很好地说明别构酶的动力学,Hill 方程是:

$$v=\frac{V_{max}\cdot[S]^h}{K_{0.5}^h+[S]^h}$$

Hill 方程与米氏方程实际上十分相似,而它们的主要差别是:首先 Hill 方程中的底物浓度[S]被提高到 h 数量级,h 被称为 Hill 系数;其次,方程底部的常数不是 K_m,而是 $K_{0.5}$,该常数也被提高了 h 数量级。$K_{0.5}$ 与 K_m 十分相似,也是指初速率为最大速率一半时候的底物浓度,但它不是米氏方程中的一部分,因此不能与 K_m 混为一谈。

Hill 系数能够反映底物协同性的程度:如果 h=1,这时的 Hill 方程就是米氏方程,也就意味着酶无底物协同性,速率对底物作图应为双曲线,$K_{0.5}$=K_m;如果 h>1,酶就具有正底物协同性,速率对底物浓度

作图呈 S 型曲线；如果 $h<1$，则意味着酶具有负底物协同性。

Quiz8 加入别构激活剂和别构抑制剂对 h 值分别有何影响？

图 9-23 显示了 h 值在 0.5 和 4 之间，速率对底物浓度的作图结果，每一种情况下酶的 V_{max} 都为 10 个单位，$K_{0.5}$ 为 4 个单位。作图的结果表明，$h=1$ 时的曲线是标准的双曲线，$h=2$ 和 4 是的曲线明显是 S 形，其中 $h=4$ 时的 S 形曲线更为明显，这意味着随着 h 值的增加，酶对底物的正协同性将提高。$h=0.5$ 的曲线显示的是负协同性，尽管形状难以和正常的双曲线区分开，但如果仔细比较，就会发现曲线在开始的时候升得很快。此外，图中所有的曲线都趋于相同的 V_{max}，各曲线也具有相同的 $K_{0.5}$ 值。

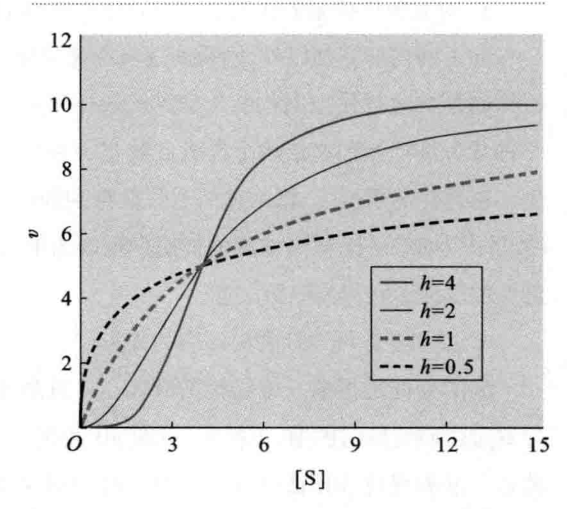
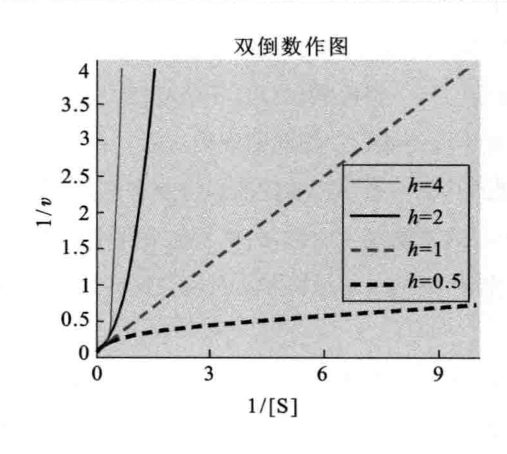


图 9-23 Hill 系数不同的酶反应速率与底物浓度的关系曲线

对米氏方程进行双倒数作图或其他线性作图法得到的是直线图，但对于不遵守米氏方程的别构酶来说，如果使用同样的线性化手段，就得不到直线图了。图 9-24 为双倒数和 Hanes 作图法得到的结果。



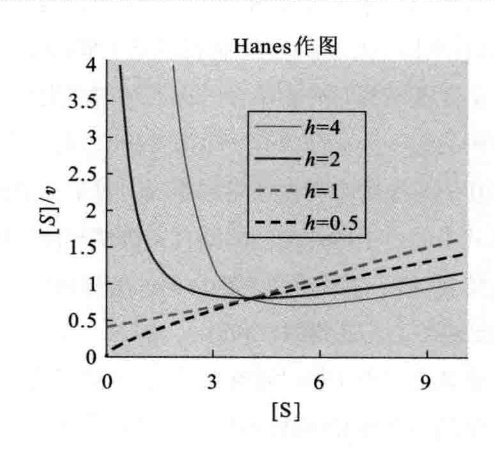


图 9-24 Hill 系数不同的酶的双倒数作图和 Hanes 作图

三、Hill 作图

虽然原来应用在米氏方程的线性化手段已不再适用于 Hill 方程，但只需对 Hill 方程稍作变换，仍然可以得到线性图。

Hill 方程可以重新整理为：$\frac{v}{V_{max}-v}=\frac{[S]^h}{K_{0.5}^h}$。如果同时对两边取对数，就得：$\lg\frac{v}{V_{max}-v}=h\cdot\lg[S]-h\cdot\lg K_{0.5}$。再以 $\lg\frac{v}{V_{max}-v}$ 对 $\lg[S]$ 作图，则得到斜率为 h、纵截距为 $-h\cdot\lg K_{0.5}$、横截距为 $\lg K_{0.5}$ 的直线（图 9-25），这种作图法就是 Hill 作图法。

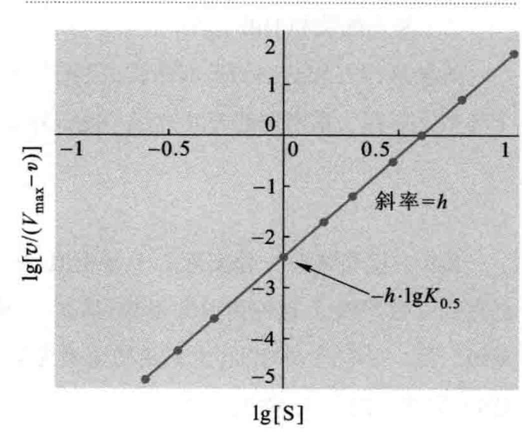


图 9-25 Hill 作图

实际上，在作图的时候，浓度很高和浓度很低的点通常偏离直线，中央部分应该是直线，这一部分可用做计算的有效数据。然而，Hill 作图遇到的难处之

一是需要事先得到 V_{max}，目前对此似乎还没有很好的解决方法。如果将米氏方程线性作图法应用到 Hill 方程中，将得不到直线。但倘若用[S]^h 代替[S]来作图，就可以得到直线，可 h 也是需要计算的，使用这样的方法计算 V_{max} 也还是不可能的！

有一种方法可以部分解决上述问题：先使用任何一种线性作图法在高底物浓度下粗估 V_{max} 值，因为这时候差不多得到的是直线。然后，使用得到的 V_{max} 进行 Hill 作图，以粗略估计出 h。随后，再用得到的 h 值，使用[S]^h 代替[S]来重新进行线性作图，这时应该会得到更好的直线。由此可以重新计算出较为准确的 V_{max}，最后再用新的 V_{max}，重新进行 Hill 作图，以得到合理的 h 值。

四、协同性的优点

既然已发现多数别构酶有底物协同性，那么协同性的出现想必会带来某些代谢上的优势。但与别构效应物相比，底物协同性的对代谢调节的重要性似乎没有那么明显。

为了更好说明底物协同性对代谢调节的重要性，需要借助于图 9–26 加以说明。图中底物浓度以 K_m 的百分数表示，并增加了上下两条横线，分别对应速率为 V_{max} 的 90% 和 10% 时的两个点。

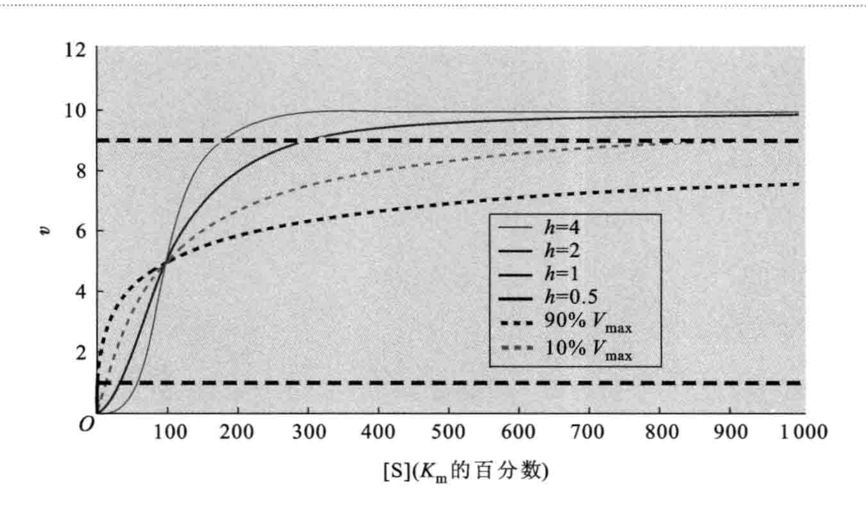


图 9–26　Hill 系数不同的酶的酶活性对底物浓度变化的敏感性

1. 正协同性的优势

仔细观察图 9–26 中的各条曲线，很容易发现，与具有正协同性的酶相比，将无协同性的米氏酶(h=1)反应速率从 V_{max} 的 10% 增加到 90% 时，需要最大的底物浓度增加。根据图中的数据，速率为 10%V_{max} 时的[S]为 K_m 的 1/9，而速率为 90%V_{max} 时的[S]为 K_m 的 9 倍，这相当于底物的浓度需要增加 81 倍。而正协同性最高的别构酶(h=4)当反应速率从 10%V_{max} 增加到 90%V_{max} 时，只需要将底物浓度提高 3 倍，正协同性低的别构酶(h=2)要达到同样的速率增长，需要提高底物浓度 9 倍。以上的结果说明，正协同效应使得酶对环境中底物浓度的变化更为敏感。这样可以让机体内某些重要的调节酶能够根据环境的变化，对代谢进行更加灵敏的调节。

2. 负协同性的优势

具有负协同效应的酶极为罕见，一个重要的例子是参与糖酵解的 3– 磷酸甘油醛脱氢酶。在图 9–26 中，要想将呈现负协同效应的别构酶(h=0.5)的反应速率从 V_{max} 的 10% 增加到 90%，需要将底物浓度提高 6 561 倍！如此大幅度的提高就意味着该酶对底物浓度的变化极度不敏感，也就保证了体内某些重要的反应不受底物浓度波动的影响，能够始终进行下去。以 3– 磷酸甘油醛脱氢酶为例，它对 NAD^+ 的结合表现为负协同性。细胞内的 NAD^+ 容易缺乏，而此酶所具有的这种性质，使得它在细胞内 NAD^+ 浓度很低的情况下，仍然具有一定的活性，从而保证了糖酵解照常进行，为细胞产生 ATP。

e9-4 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/enzyme-kinetics.html（完全免费的医学生物化学网站，有关酶动力学的内容）
2. http://chemwiki.ucdavis.edu/Biological_Chemistry/Catalysts/Enzymatic_Kinetics（美国加利福尼亚大学戴维斯分校一个网站，有关酶动力学的内容）
3. http://en.wikipedia.org/wiki/Enzyme_kinetics（维基百科有关酶动力学的内容）

参考文献：

1. Bennet W S, Steitz T A. Glucose-induced conformational change in yeast hexokinase. *Proc Natl Acad Sci*, 75:4 848–4 852, 1998.
2. Cleland W W. Determining the chemical mechanisms of enzyme catalyzed reactions by kinetic studies. *Adv Enzymol*, 43:273–387, 1977.
3. Rando R R. Mechanisms of action of naturally occurring irreversible enzyme inhibitors. *Acc Chem Res*, 8:281–288, 1975.

第十章　酶的催化机理

酶最大的神奇之处在于其高效的催化能力,那么是什么导致神奇出现的呢?为此,科学家对一系列酶的催化机理进行了大量的研究,并已经取得了丰硕的成果,像对某些酶作用机理(例如丝氨酸蛋白酶)的认识早已十分清楚。尽管如此,人们仍然对很多酶的催化机理了解得不多,甚至是一无所知。可以说,揭示出一种酶的催化机理是酶学研究中最具挑战性的工作。

本章将首先简单介绍研究酶催化机理的主要方法,然后重点介绍酶催化的"过渡态稳定"学说及稳定过渡态的若干机制,最后以几种常见的酶为实例,详细介绍它们的催化机制。

第一节　酶催化机理研究的主要方法

要揭示一种酶的催化机理,首先应该选择正确的研究方法。用来研究酶催化机理的主要方法有:

1. 确定酶的一级结构,找出保守性的氨基酸残基

在得到不同生物来源的同一种酶的一级结构以后,将它们进行比对,找出高度保守的氨基酸残基,有可能获得一些与催化有关的氨基酸残基的信息。显然,在一种酶的进化过程中,保守性越高的氨基酸残基对酶的三维构象和功能贡献就越大,这些氨基酸残基很可能与酶的催化机理直接相关。例如,在对各种丝氨酸蛋白酶的一级结构进行比较后发现,有3个氨基酸残基即H57、D102和S195是不变的。事实证明,这3个残基直接参与肽键的水解。

2. 确定酶的三维结构

酶的功能与其三维构象直接相关联,如果能够得到一种酶的三维结构的信息,就会为其催化机制的研究提供最直接的线索。获得酶三维构象的方法主要有X射线晶体衍射和NMR。要想利用前一种方法获取一种酶的三维构象,那么必须先得到它的晶体,才能对其进行X射线衍射分析,并比较有无底物或抑制剂时的构象变化,从而获取酶分子特别是活性中心的构象。不幸的是,获取一个酶的晶体并非易事。

3. 定点突变

分子生物学技术的发展使蛋白质的定点突变已变得十分容易,实际上它已成为实验室的一项常规技术。理论上,只要得到一个酶的基因,就可以对它的一级结构进行"随心所欲"的突变。如果在其他研究的基础上,对一种酶分子上的某个或某些氨基酸残基(特别是活性中心附近)进行有的放矢的突变,就会对其作用机理的研究提供更有价值的信息。例如,有人将胰蛋白酶活性中心裂缝底部的D189进行突变,使其成为不带电荷或带正电荷的氨基酸残基,结果发现胰蛋白酶活性会完全丧失。联想到胰蛋白酶的底物专一性(只水解碱性氨基酸羧基与其他氨基酸氨基形成的肽键),人们便意识到,胰蛋白酶在催化的时候,是通过D189带负电荷的侧链与底物上的碱性氨基酸残基带正电荷的侧链结合的。

4. 动力学分析

动力学分析也可以为研究酶的催化机理提供一些有用的数据,像胰凝乳蛋白酶的"暴发动力学"(burst kinetics)性质就能表明,由它催化的水解反应分为快慢两个阶段(参看本章"丝氨酸蛋白酶的作用机制")。在动力学研究中,需要特别注意pH以及特定抑制剂对酶活性的影响。以pH为例,由于溶液的pH能影响到活性中心侧链可解离基团的解离状态,我们便可以通过研究解离状态和反应速率之

间的关系，来推断出哪些氨基酸残基与酶催化有关。同样，利用各种酶抑制剂作用的动力学也可以提供一些重要的信息。

5. 化学修饰

化学修饰不仅可用来改变酶的性能(参看第十三章“酶的应用及研究方法”)，还能为酶的作用机理的研究提供一些有价值的信息。例如，科学家在研究胰凝乳蛋白酶作用机制的时候，就使用 DIPF 和 TPCK 分别确定了 Ser 和 His 在催化中的作用。

如果使用的修饰试剂是双功能试剂，就有助于确定一种酶分子在三维结构上相邻的基团；如果在进行化学修饰的同时，利用底物或竞争性抑制剂对活性中心进行保护，就可以确定活性中心内的氨基酸残基的性质。例如，假定一个酶分子共有 6 个 Cys 残基，但无法确定它的活性中心是否有或者有几个 Cys 残基。这时候可以分别在有无过量底物或竞争性抑制剂的存在下，利用碘代乙酸修饰酶，并对两种条件下的修饰产物进行比较。显然，在有底物保护的情况下，不能被修饰的 Cys 残基就在活性中心。此外，还可以先在过量底物或竞争性抑制剂存在的条件下，使用非同位素标记的碘代乙酸修饰酶。然后，利用透析或凝胶过滤除去底物或抑制剂，并使用同位素标记的碘代乙酸重新修饰酶，那么能被同位素修饰的 Cys 残基就应该位于活性中心。

Quiz1 为什么酶活性中心的 Cys 残基更容易与碘代乙酸起反应？

6. 计算机模拟

尽管 X 射线晶体衍射和 NMR 等技术对于研究酶的三维结构十分重要，可是它们有时并不能提供一种酶催化机制的详细信息。然而，如果在用实验的手段进行研究的同时，能结合计算机模拟，这就为深入了解酶的催化机制提供极为宝贵的信息。通过计算机模拟，可以对酶催化反应中形成的过渡态、亚稳态中间物、酶-底物(ES)或者酶-产物(EP)复合物的结构和性质进行直接的预测和鉴定，而且可以在原子水平上提供更为精细的分子结构，并对反应的能量学进行精确的计算和分析，在此基础上有可能搞清楚每一个基团对催化的贡献。

第二节 过渡态稳定学说

根据过渡态理论(transition state theory)，在任何一个化学反应系统中，反应物需要到达一个特定的高能状态以后才能发生反应。这种不稳定的高能状态被称为过渡态(transition state)。过渡态一般在形状上既不同于反应物，又不同于产物，而是介于两者之间的一种不稳定的结构状态，这时候旧的化学键在减弱，新的化学键开始形成。过渡态存留的时间极短，只有 10^{-14}~10^{-13} s。要达到过渡态，反应物必须具有足够的能量以克服势能障碍，即活化能(activation energy)。然而，一个反应系统中各反应物分子具有不同的能量，只有某些反应物具有足够的能量去进行反应。

需要特别注意的是，活化能($\Delta G^{\ominus}$)并不等同于反应的总自由能变化(ΔG)。两者的差别在于：ΔG 与反应平衡时底物浓度和产物浓度有关，而 $\Delta G^{\ominus}$ 与反应速率有关。显然，$\Delta G^{\ominus}$ 越低，达到过渡态的反应物分子就越多，反应发生得就越快！提高温度或加入催化剂可以促进更多的反应物达到过渡态，从而提高反应速率。

一、过渡态稳定学说

酶之所以能够催化反应，是因为它能降低反应的活化能。实际上，活化能的小幅度下降可导致反应速率大幅度的提升。以过氧化氢分解成氧气和水的反应为例，在没有任何催化剂时反应的活化能为 76 kJ/mol，而在有过氧化氢酶的情况下，活化能仅仅降到 30 kJ/mol，但反应速率却提高了 10^8 倍，这样的提升相当于将几年才能完成的反应缩短到几秒钟。那么，酶是如何降低一个反应的活化能的呢？

首先以一个虚拟的催化铁丝断裂的“铁丝酶”为例加以说明：如图 10–1 所示，铁丝要发生断裂，首先必须克服能障，进入弯曲发热、似断非断的过渡态，然后才有可能一分为二变成产物。如果没有酶的催化，铁丝断裂的反应式就是 S → $S^{\ominus}$（过渡态）→ P，其活化能是 $\Delta G_N^{\ominus}=G_S^{\ominus}-G_S$；如果有酶的催化，反应的路径就变成了 E+S → ES → $ES^{\ominus}$（过渡态）→ E+P，这时反应的活化能 $\Delta G_E^{\ominus}=G_{ES}^{\ominus}-G_{ES}$。显然只有 $\Delta G_E^{\ominus}<\Delta G_N^{\ominus}$，酶才能催化。然而，图 10–1（2）显示的酶是催化不了铁丝的断裂的，因为它的活性中心与基态的铁丝互补，而不是与过渡态的铁丝互补。无疑铁丝进入这样的活性中心很容易，但一旦进入，铁丝会以较高的亲和力进入一种十分稳定的 ES 状态，这时候的铁丝实际上被活性中心冻结住了，要进入过渡态 $ES^{\ominus}$需要克服更大的活化能（$\Delta G_E^{\ominus}>\Delta G_N^{\ominus}$），因而反而不利于反应。而图 10–1（3）显示的酶就能有效地催化，因为它的活性中心与铁丝的过渡态互补。当铁丝进入这样的活性中心以后，受活性中心各种反应性基团的作用，不得不发生形变，随后很容易进入过渡态，因为酶与过渡态的亲和性更高。在这种情况下的 $\Delta G_E^{\ominus}<\Delta G_N^{\ominus}$，即活化能降低了，因此酶发生了催化。由此可见，一种酶要行使催化，它必须能够通过某种方式稳定反应的过渡态，而不是稳定基态，即酶与过渡态的亲和力必须要高于它与底物的亲和力。酶的这种催化反应的机制被称为“过渡态稳定”学说。

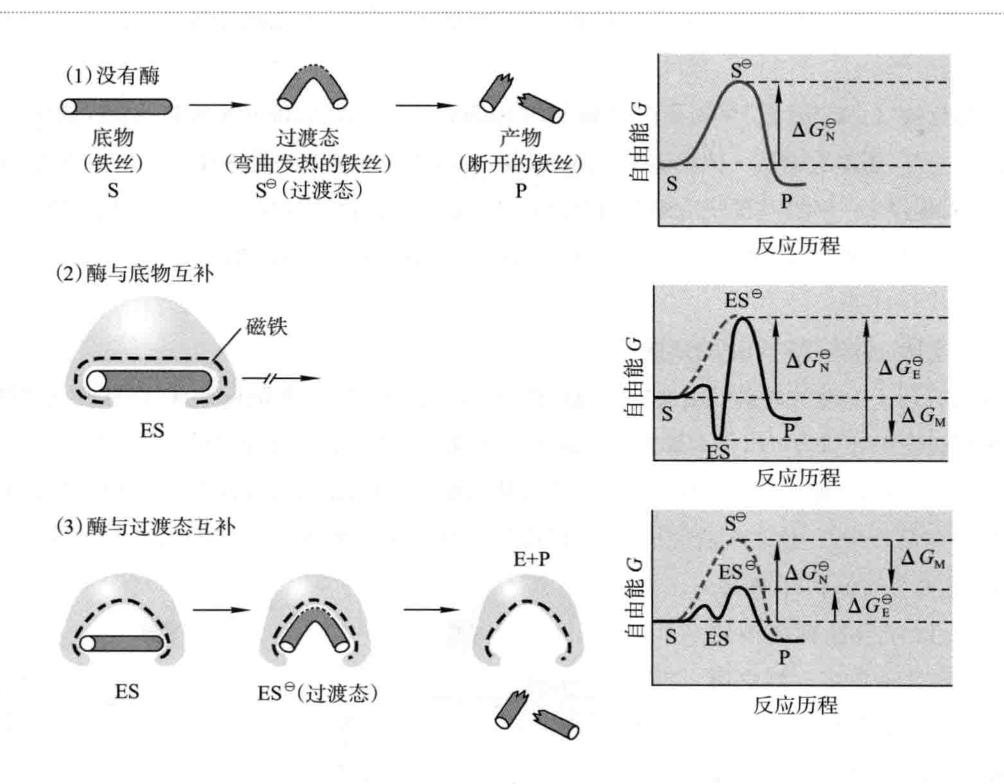
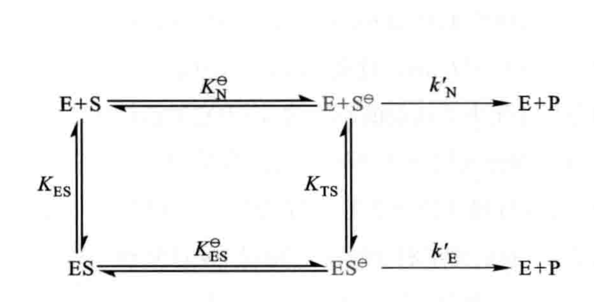


图 10–1 “铁丝酶”催化反应的过渡态稳定机制（Nelson 等，2008）

再以一个真正的单底物 – 单产物反应（S → P）为例作进一步说明（图 10–2）：如果没有酶，S 需要克服能障变成过渡态 $S^{\ominus}$以后才能转变成 P；如果有酶的催化，就首先通过典型的分子碰撞，酶与底物形成可逆的复合物 ES。在典型的实验条件下，平衡有利于 ES 的形成，这是由于一对典型 ES 形成的结合能为 –50.2~–12.5 kJ/mol。ES 形成的结合能有一部分被用来驱动过渡态 $ES^{\ominus}$的形成。由于形成 $ES^{\ominus}$需要的能量（$\Delta G_E^{\ominus}$）比单独形成 $S^{\ominus}$所需要的能量（$\Delta G_N^{\ominus}$）低，因此反应的总活化能显著降低，酶促反应的速率也就随之大增。但是，酶并不能改变反应的总 ΔG，即它不能影响反应的平衡常数。

酶与底物结合而得到的结合能主要由结构的互补性（如较好的三维结构的契合以及合适的非共价的离子键和氢键作用力）决定，但并不是结合能越大越好。许多非酶蛋白质与它们各自的配体结合产生的结合能非常大，却没有催化效应。例如，抗体 – 抗原的解离常数接近 10^{-8} mol/L，其结合能是 –46 kJ/mol，而 ATP 与肌球蛋白结合的解离常数是 10^{-13} mol/L，它们的结合能是 –75 kJ/mol，生物素



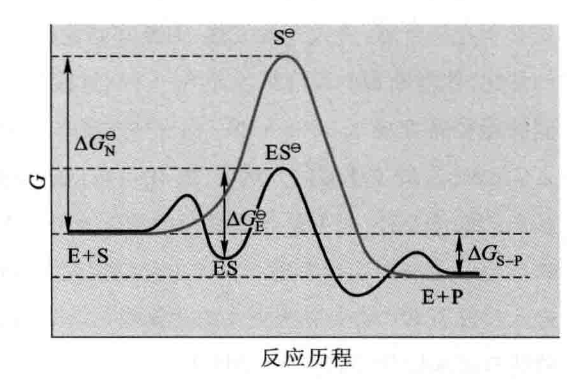


图 10-2 酶促反应和非酶促反应的反应历程及活化能与自由能的变化

与抗生物素蛋白(avidin)结合的解离常数为 10^{-15} mol/L,它们的结合能为 −86 kJ/mol。从这些数据可以看出,结合能越大,蛋白质与配体的亲和力越高。如果一种酶与底物结合产生的结合能太大,两者的亲和力就过强,于是形成过渡态 ES$^{\ominus}$就需要更大的能量用于克服 ES 之间过强的亲和力,这时反而不利于催化。

酶催化的“过渡态稳定”学说早在 1946 年由 Pauling 提出。R. Wolfenden 曾打过这样的比喻,酶就像助产士一样,帮助底物从一种亚稳态(metastable state)艰难地进入到另一种亚稳态。 总之,酶在催化的时候,提供了一种替代的反应路径,将反应分割成若干步,因其与反应过渡态的结合比和底物的结合更牢、更好,稳定了反应的过渡态,从而降低活化能,使得反应速率加快。

二、支持过渡态稳定学说的证据

现在有两个关键的证据可用来支持“过渡态稳定”学说。第一个证据是,根据过渡态的结构设计出来的过渡态类似物可以作为酶的强抑制剂,其抑制效果要比竞争性抑制剂强得多。因为要是酶的确能稳定过渡态,那么人工设计的过渡态类似物只要遇到酶,就会与活性中心紧密地结合,并牢牢地卡住酶的活性中心,使之无法完成反应,即成为了酶的强抑制剂。迄今为止,人们已成功得到多种酶的过渡态类似物抑制剂,它们的 K_i 比竞争性抑制剂要小 1 000 倍甚至更多。其中第一例与脯氨酸消旋酶(proline racemase)有关。根据 Pro 过渡态设计得到的两种过渡态类似物——吡咯 −2− 羧酸(pyrrole-2-carboxylate)和 Δ−1− 吡咯碱 −2− 羧酸(Δ−1-pyrroline-2-carboxylate),它们都是脯氨酸消旋酶的强抑制剂(图 10-3)。

图 10-3 脯氨酸消旋酶的过渡态及其过渡态类似物抑制剂

支持“过渡态稳定”学说的第二个证据是,利用过渡态类似物作为抗原或半抗原,去免疫动物,由此产生的抗体可能具备类似酶的催化作用。

早在 1969 年,W. P. Jencks 在他编著的名为《化学和酶学中的催化》(*Catalysis in Chemistry and Enzymology*)一书中提到:“如果酶如描述的那样通过与过渡态更紧密地结合,通过与这种状态具有最大的相互作用而稳定一个反应的过渡态,那么,合成一种酶的途径是制备与一个给定反应的过渡态相

似的半抗原基团的抗体，这样的抗体的结合位点应该与过渡态互补，并且迫使结合的底物接近过渡态而促使反应加速”。

既然动物的免疫系统能够制造大约 10^9 种以上不同类型的抗体，那么就可能制造出能合成的任何分子的抗体。但问题是如何选择和合成正确的过渡态类似物，以及如何制造出大量的合适抗体。第二个问题随着杂交瘤技术的发展得到了解决，而第一个问题的解决取决于对反应本身的性质、复杂性以及对反应机理的了解。

1986 年 12 月，有两个独立的研究小组在 *Science* 上同时发表了成功得到抗体酶的论文。他们的论文都是建立在一个事实之上：磷酸酯在磷原子上呈四面体状，这与羧酸酯水解过程中经历的过渡态相似（图 10–4）。首先获得抗体酶的是 Pollack S. J.，他以羧酸二酯水解反应的过渡态类似物——对硝基苯酚磷酸胆碱酯，作半抗原诱导产生单克隆抗体。经过筛选，从中找到一株可催化水解反应

图 10–4　过渡态类似物的预测和设计

框 10–1　生化制药——水解可卡因药物的研制

许多让人上瘾的毒品，如可卡因，通过与神经元上特定受体的结合，模拟某种神经递质的作用而作用。当一个人沉迷于这种药物，治疗成瘾的常规方法是使用一种化合物阻断受体，让毒品结合不了。

可卡因成瘾一直难以治疗，主要是由于其独特的作用方式。可卡因在体内可阻断神经递质多巴胺的重吸收。结果使得多巴胺在系统中存留更长的时间，去过度刺激神经细胞，不断在大脑中产生奖赏信号，而导致成瘾。因此使用阻断受体的药物来治疗可卡因成瘾是无效的，反而可能会使多巴胺去除更困难。

可卡因的化学结构上含有一个酯键，因此理论上是可以水解的。但在人体内并没有一个天然的酯酶能催化它的水解（图 10–5）。然而，根据酶催化的过渡态稳定学说，如果可卡因发生水解，就必须经过一个过渡态。这样完全可以根据筛选抗体酶的思路，制造出专门催化可卡因水解的催化性抗体。

图 10–5　可卡因水解的过渡态及其类似物的结构

令人兴奋的是，已有科学家成功制备出了水解可卡因的抗体酶，当让可卡因成瘾患者服用这样的抗体酶以后，可卡因在他们的体内被成功地水解成 2 个无害的产物——苯甲酸和芽子碱甲酯。而一旦可卡因降解，就不能再阻止多巴胺的重吸收，也就没有神经刺激时间的延长发生。于是，药物成瘾的影响随着时间的推移而消失。

的 MOPC 167,能使速率加快 12 000 倍。该抗体催化反应的动力学行为满足米氏方程,并具有底物特异性及 pH 依赖性等酶促反应的特征。

酶过渡态类似物抑制剂和抗体酶的存在为 Pauling 的假说提供了强有力的证据,可以说,要是酶与过渡态的亲和性还不及与底物的亲和性,那么酶充其量只是一种底物结合蛋白,而不会是一种催化剂。

第三节 过渡态稳定的化学机制

与酶催化相关联的过渡态稳定是酶活性中心的结构和反应性,以及活性中心与结合的底物之间相互作用的必然结果。酶在行使催化功能的时候,会充分利用各种化学机制来实现过渡态的稳定并由此加速反应。这些机制归纳起来主要有 6 种:邻近定向效应(proximity and orientation)、广义的酸碱催化(general acid/base catalysis)、静电催化(electrostatic catalysis);金属催化(metal ion catalysis)、共价催化(covalent catalysis)和底物形变。

一、邻近定向效应

邻近定向效应是指两种或两种以上的底物(双底物居多)同时结合在酶活性中心上,相互靠近(邻近),并采取正确的空间取向(定向)。这样可以大大提高底物的有效浓度,使分子间反应近似分子内反应,从而加快反应速率。

底物与活性中心的结合不仅使底物与酶催化基团或其他底物接触,还强行"冻结"了底物的某些化学键的平动和转动,促使它们采取正确的取向,也有利于键的形成。

为了便于说明,让我们假定一个双分子反应,在底物 A 和底物 B 之间形成共价键,产生化合物 A–B,即 A + B → A–B。

如果这个反应完全在溶液中发生,就需要满足以下条件:① A 和 B 通过受扩散限制的碰撞相遇,相互间还需要采取正确的取向;②发生脱溶剂化变化,以使分子轨道能相互作用;③克服范德华斥力;④电子轨道发生变化以进入过渡态。显然,在溶液中的反应速率是由两种底物有效碰撞的概率来决定。虽然提高温度和底物浓度可以大幅度地增加底物间有效碰撞的机会,但是对反应速率的提升是十分有限的。

如果这个反应是酶促反应,那么底物与酶活性中心的结合是反应发生的前提。当底物被隔离在酶的活性中心内,底物的有效浓度相对于在溶液中的浓度就大大增加了。其次,酶活性中心的结构被预先设计得正好可与底物按特异的方向结合,非常有利于反应的进行。在大多数双分子反应中,两个底物之间必须采取合适的取向,才能进入过渡态。然而,在溶液中,每一种底物都存在一群旋转异构体(rotomer),它们的存在能阻滞反应速率。如果两种底物被锁定在酶的活性中心,则是另外一种情景,酶可以轻而易举地克服那些底物达到过渡态的障碍。当然,底物在活性中心遭受的在空间和取向上的严重限制减少了反应的熵。但总的来说,酶与底物的结合能会弥补熵的减少。

二、广义的酸碱催化

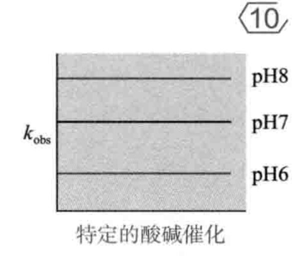
广义的酸碱催化是指水分子以外的分子作为质子供体或受体参与催化,这种机制几乎参与了所有酶的催化。与广义酸碱催化有关的化学基团包括:具有 3 个 pK_a 的氨基酸(如 Cys、Ser、Asp、Glu、Lys 和 His)的 R 基团,以及肽链两端游离的 α– 氨基和 α– 羧基。对于核酶而言,参与广义酸碱催化的基团包括核糖的 2′–OH、腺嘌呤的 N1 和胞嘧啶的 N3。需要特别注意的是,由于微环境的变化,这些基团在一个酶分子上真实的 pK_a 与它们在游离的氨基酸或核苷酸上的 pK_a 差别可能很大。

这些起催化作用的基团可以提供质子或者接受质子来稳定过渡态上的电荷，同时还具有激活亲核基团、亲电基团或稳定离去基团的功能。如果一个侧链基团的 pK_a 接近 7，该侧链基团就可能是最有效的广义酸碱催化剂。His 残基的咪唑基就是这样的基团，因此它是很多酶的催化残基。

广义的酸碱催化不可以和特定的酸碱催化混为一谈，两者的主要差别在于：①后者是 H^+ 或 OH^- 加速反应，前者是酶分子上的酸性和碱性基团催化反应；②缓冲溶液的浓度对它们的影响不一样（图 10-6），特定的酸碱催化反应速率只取决于 pH，与缓冲溶液浓度没有关系，而就广义的酸碱催化而言，缓冲溶液有助于通过提供或夺取质子来稳定过渡态，因此反应速率不仅取决于底物和酶的质子化状态，还取决于缓冲溶液的浓度。

在生物体内，各种酯的水解经常使用广义酸碱催化。如图 10-7 所示，在酯水解中，广义酸（HA）作为酯羰基氧的部分质子供体，增强了酯分子上部分负电荷的过渡态的稳定性；或者，广义碱（B）充当质子受体来稳定过渡态。有时，为了进一步提高催化效率，许多酶会同时利用广义酸催化和广义碱催化。

以核糖核酸酶 A 为例（参看图 10-8），其活性中心的 His 12 和 His119 分别充当广义碱和广义酸行使催化。在反应中，His12 作为质子受体，从 3′－核苷酸的 2′－羟基抽取 1 个质子，而 His119 作为质子



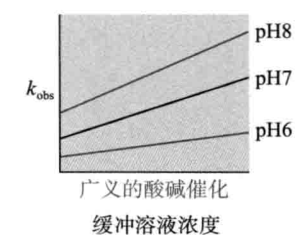


图 10-6 pH 和缓冲溶液浓度对特定酸碱催化和广义酸碱催化的影响

图 10-7 广义的酸催化和碱催化

图 10-8 核糖核酸酶 A 的广义酸碱催化

供体，提供 1 个质子给核苷酸的 5′- 羟基。其净结果是质子从 His119 传到 His12。随后，水取代释放的核苷酸，这时 His12 和 His119 的碱酸角色正好颠倒，原来的 His 质子化状态得到恢复。

三、静电催化

对于很多酶而言，活性中心电荷的分布可用来稳定酶促反应的过渡态，酶使用自身带电基团（有时是带部分电荷的基团），去中和一个反应过渡态形成时所产生的相反电荷而进行的催化，被称为静电催化。计算机模拟的研究表明，静电效应可能是对酶催化机制中贡献最大的因素。由于酶的活性中心有时可以提供比水更强的极性环境，因此带电的过渡态可以被固定的偶极更好地稳定。这完全不同于水环境中的过渡态稳定，因为在水环境中的稳定需要消耗能量去重排水分子。

酶分子经常提供带电基团的有 Lys、Arg、Asp 和 Glu 的侧链以及金属离子。有时，酶通过与底物的静电作用将底物引入到活性中心。

四、金属催化

近三分之一已知酶的活性需要金属离子的存在，这些酶分为两类，一类为金属酶（metalloenzyme），另一类为金属激活酶（metal-activated enzyme）。前者含有紧密结合的金属离子，多数为过渡金属，如 Fe^{2+}、Fe^{3+}、Cu^{2+}、Zn^{2+}、Mn^{2+} 或 Co^{3+}，后者与溶液中的金属离子松散地结合，通常是碱金属或碱土金属，例如 Na^{+}、K^{+}、Mg^{2+} 或 Ca^{2+}。

金属离子参与的催化被称为金属催化。金属离子以 5 种方式参与催化：①作为路易斯酸（Lewis acid）接受电子，使亲核基团或亲核分子（如水）的亲核性更强；②与带负电荷的底物结合，屏蔽负电荷，促进底物在反应中正确定向，例如，所有的激酶都需要 Mg^{2+}，Mg^{2+} 所起的作用是屏蔽底物 ATP 分子上所带的高度负电荷，并对 ATP 起定向的作用；③参与静电催化，稳定带有负电荷的过渡态；④通过价态的可逆变化，作为电子受体或电子供体参与氧化还原反应；⑤本身是酶结构的一部分。

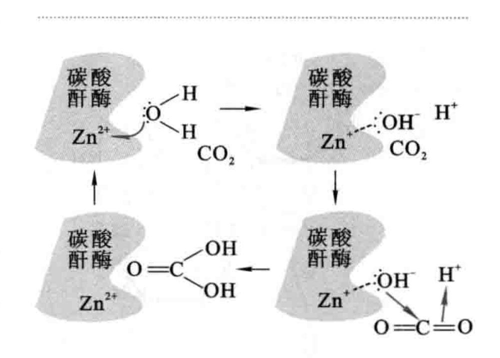


图 10-9　碳酸酐酶的金属催化机制

以碳酸酐酶为例（图 10-9）：从水分子上解离下来的 OH^{-}，被酶活性中心上与咪唑基结合的 Zn^{2+} 接受以后，亲核性大增，于是能更有效地进攻 CO_2 上的 C 原子，并最终产生 HCO_3^{-}。

再以四膜虫体内催化 26S rRNA 前体剪接的 Ⅰ 组内含子为例，这种核酶利用金属 Mg^{2+} 直接稳定带负电荷的过渡态（参看第十一章“核酶的结构与功能”）。

五、共价催化

共价催化需要酶在催化过程中，与底物暂时形成不稳定的共价中间物。共价中间物的形成完全改变了反应的路径，新的路径所需要的活化能大大低于没有酶催化时的活化能，这有利于克服活化能能障。行使共价催化的酶，已进化到能将这类困难的反应分两步进行，一是共价中间物的形成，二是共价中间物的断裂，而不直接催化单个反应。因此，反应的限速步骤要么是共价中间物的形成，要么是共价中间物的断裂。例如，胰凝乳蛋白酶催化的水解反应的限速步骤就在第二步。

酶与底物形成共价中间物的手段主要是亲核催化（nucleophilic catalysis），也可以是亲电催化（electrophilic catalysis）。前者是酶分子上的亲核基团对底物做亲核进攻而引发反应，后者则是由酶分子上的亲电基团对底物做亲电进攻而启动反应。

酶分子上能进行亲核进攻的亲核基团有：3 种羟基氨基酸的羟基、Cys 的巯基、Lys 的 ε- 氨基、两种酸性氨基酸的羧基和作为辅酶的焦磷酸硫胺素（TPP）（表 10-1）。

表 10-1 几种酶的共价催化

亲核基团	实例	共价中间物
Ser(—OH)	丝氨酸蛋白酶	脂酰化酶
Cys(—SH)	半胱氨酸蛋白酶	脂酰化酶
Asp(—COO—)	Na^+/K^+-ATP 酶	磷酸化酶
Lys(ε-NH_2)	乙酰乙酸脱羧酶和第一类醛缩酶	希夫碱
His(咪唑基)	磷酸甘油酸变位酶	磷酸化酶
Tyr(—OH)	谷氨酰胺合成酶	腺苷酸化酶
TPP	丙酮酸脱羧酶和转酮酶	羟乙基化酶

亲核催化涉及酶分子上的亲核基团提供电子给底物,先形成带部分共价键的过渡态中间物,再形成真正的共价中间物。其速率取决于进攻性的亲核基团的亲核性(供电子能力)和底物对进攻基团的敏感性。以图 10-11 为例,好的亲核反应应该是:Y 是比 X 更好的离开基团(leaving group),X 是比 Z 更好的进攻基团,共价中间物的反应性应该比底物强。

在各种亲核催化中,Lys 残基上的 ε- 氨基是很好的亲核基团,以乙酰乙酸脱羧酶为例(图 10-11),其活性中心的一个 Lys 残基的 ε- 氨基是一个亲核进攻基团(相当于图 10-10 上的 X),与乙酰乙酸上的 β- 酮基形成希夫碱(Schiff base)。希夫碱上质子化的 N 原子充当一个有效的"电子穴"(electron sinks)对相邻的碳负离子起稳定作用,而羧基(相当于 Y)又是一个很好的离开基团(脱羧)。

图 10-10 亲核共价修饰反应

图 10-11 乙酰乙酸脱羧酶的亲核催化

希夫碱是常见的酶底物共价中间物,以这种机制进行催化的酶通常受到硼氢化钠(sodium borohydride)的不可逆抑制,该金属还原剂能够将希夫碱还原成不能断裂的中间物,致使产物无法释放。反过来,硼氢化钠也经常被用来验证希夫碱是否参与一种酶的共价催化。

在亲电催化中,共价中间物的形成源于酶分子上的亲电基团对底物分子上富电子基团的亲电进攻。由于酶分子上的氨基酸侧链并不能提供有效的亲电基团,因此使用亲电催化的酶只能依靠其缺乏电子的辅酶或金属离子,例如磷酸吡哆醛。

Quiz2 为什么单独的辅酶行使不了催化?

使用磷酸吡哆醛作为辅酶的酶有转氨酶、消旋酶、脱羧酶和糖原磷酸化酶等,但糖原磷酸化酶并不是利用磷酸吡哆醛进行亲电催化,而是利用其中的磷酸基团辅助催化(参看第二十七章"糖原代谢"),这与转氨酶、消旋酶和脱羧酶完全不同。这里只介绍使用磷酸吡哆醛的亲电催化。

磷酸吡哆醛在亲电催化中的实际功能是充当电子穴,稳定催化过程中形成的碳负离子。由于阳

离子亚胺(cationic imine)与芳香杂环－吡啶环相连，电荷分布得以分散，使得磷酸吡哆醛成为一个非常有效的亲电催化剂。利用吡哆醛亲电催化的酶促反应都由3步组成(图10-12)：①阳离子亚胺的形成；②经过碳负离子中间物的化学变化；③亚胺产物的水解。在各种与氨基酸相关的反应(包括转氨、消旋、脱羧和侧链互变反应)中，磷酸吡哆醛都是与α－氨基酸反应，去除α－氢，产生希夫碱中间物。以转氨反应为例，在α－氢消除以后，质子供给磷酸吡哆醛的羰基碳，形成一种α－酮酸－吡哆胺式希夫碱。随后，希夫碱水解，产生游离的α－酮酸和磷酸吡哆胺。紧接着，另一种α－酮酸与磷酸吡哆胺之间形成亚胺，而质子朝相反的方向转移，从而产生新的氨基酸，同时再生出吡哆醛而完成催化循环。

图 10-12 磷酸吡哆醛的亲电催化

六、底物形变

底物形变是诱导契合产生的主要效应。酶对底物的诱导导致酶的活性中心与过渡态的亲和力高于它与底物的亲和力，当酶与底物相遇时，酶分子诱导底物分子内敏感键更加敏感，产生“电子张力”发生形变，从而更接近它的过渡态，由此降低了反应的活化能并有利于催化反应的发生。然而，事实上形变诱发更多的是对基态底物的去稳定效应，而不是对过渡态的稳定效应。此外，除了底物发生形变以外，酶本身也可能发生形变，从而导致活性中心的某些直接参与催化的氨基酸残基被激活。前面提到的虚拟的“铁丝酶”也许最能说明这种催化机制，现实中的溶菌酶也利用这种方式进行催化。在溶菌酶催化的时候，与其活性中心结合的六碳糖在溶菌酶的“威胁利诱”下，从椅式构象变成半椅式构象而发生形变，周围的糖苷键也更容易发生断裂(见后关于溶菌酶催化机制的内容)。

第四节 几种常见酶的结构与功能

一、蛋白酶

蛋白酶是催化肽键水解的一类酶的总称。虽然肽键的水解在热力学上是十分有利的，但如果没有蛋白酶的催化，一个肽键在中性pH和25℃条件下需要300~600年的时间才能完成水解。

根据活性中心催化基团的性质，蛋白酶可以分为4类：①丝氨酸蛋白酶，此类蛋白酶的催化基团包括1个不可缺少的Ser残基，DIPF是它们的不可逆性抑制剂；②巯基蛋白酶（thiol protease），其催化基团包括1个Cys的巯基，巯基试剂（如碘代乙酸）是它们的不可逆性抑制剂；③天冬氨酸蛋白酶（aspartic protease），其催化基团包括两个重要的Asp，在偏碱性的pH下无活性；④金属蛋白酶（metalloprotease），此类蛋白酶在活性中心结合金属离子（如Zn^{2+}），金属螯合剂EDTA或邻菲罗啉（*o*-phenanthroline）可导致其活性的丧失。

所有蛋白酶在催化过程中都要经历四面体的过渡态，过渡态的形成涉及一个亲核基团进攻肽键的羰基碳。

（一）丝氨酸蛋白酶和巯基蛋白酶

这两类蛋白酶在催化反应中，先后就有两次亲核进攻，其中第一次亲核进攻的基团分别是Ser残基的羟基和Cys残基的巯基，第二次都是水分子。催化机理中包括广义酸碱催化和共价催化等，脂酰化酶是酶与底物的共价中间物。

1. 丝氨酸蛋白酶

这一类蛋白酶以一个特定的Ser残基作为必需的催化基团，属于该家族成员的有胰蛋白酶、胰凝乳蛋白酶、弹性蛋白酶、枯草杆菌蛋白酶（subtilisin）、激肽释放酶原（kallikrein）、凝血酶（thrombin）、纤溶酶（plasmin）和组织型纤溶酶原激活剂（tissue plasmin activator，tPA）。其中前3种酶为动物体内的消化酶，最后两种为调节蛋白酶，参与与血凝或溶栓相关的级联放大过程。激肽释放酶原也是一种调节蛋白酶，它通过切除激素原和生长因子的"原序列"来激活它们。枯草杆菌蛋白酶是由枯草杆菌分泌的消化酶。

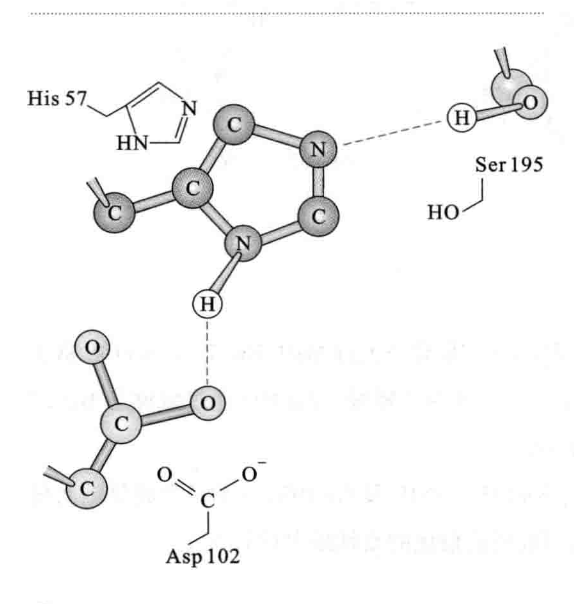


图 10-13 丝氨酸蛋白酶的催化三元体

参与消化过程的各种丝氨酸蛋白酶对底物具有不同的专一性。专一性都与肽键羰基一侧的氨基酸残基的侧链性质有关。决定专一性的是酶活性中心的立体化学：胰蛋白酶具有很深的底物结合口袋，而口袋的底部又是一个带正电荷的Asp，故特别适合长的碱性氨基酸侧链（Lys和Arg）的结合；胰凝乳蛋白酶的底物结合口袋没有胰蛋白酶深，但比它宽，而且分布在口袋壁上的是疏水氨基酸，因此特别适合与芳香族氨基酸的结合；弹性蛋白酶的底物结合口袋非常浅，最适合结合侧链较小的氨基酸残基。

Quiz3 与胰凝乳蛋白酶相比，胰蛋白酶很容易发生自我消化，为什么？

丝氨酸蛋白酶的催化机制属于共价催化和广义酸碱催化的混合体，由3个不变的氨基酸残基构成的催化三元体（Ser、His和Asp）起主导作用（图10-13），这3个氨基酸残基的任何一种若发生突变或被化学修饰（例如，Ser被DIPF修饰），均会导致酶活性的丧失。

虽然各种丝氨酸蛋白酶具有高度的同源性，但构成催化三元体的3个氨基酸残基在肽链中的位置并不完全相同。然而，酶学家已达成共识，将这3个氨基酸残基总是进行相同的编号，即His57、Asp102和Ser195。它们在催化中的具体功能是：① Ser195充当进攻底物的亲核基团；② His57作为广义的酸碱催化剂；③ Asp102的功能仅仅是定向His57，影响His的pK_a，改变其酸碱解离性质。

Quiz4 如果你发现一种新的蛋白酶，你如何确定它是否属于丝氨酸蛋白酶家族？

以胰凝乳蛋白酶为例，其除了具有蛋白酶活性以外，还具有缓慢的酯酶活性。例如，对硝基苯酚乙酸酯就是其酯酶活性的底物，它特别适合研究胰凝乳蛋白酶的作用机理。

胰凝乳蛋白酶的催化经历了脂酰化酶共价中间物，全部反应分为5步（图10-14）：

(1) 底物与活性中心的Ser195和Gly193残基通过氢键形成ES，疏水口袋正好能够容纳芳香环。

(2) 亲核的Ser195残基进攻羰基碳，形成第一个四面体形的过渡态中间物，其中有一个氧以氧阴

图 10-14 胰凝乳蛋白酶催化的蛋白水解反应的全过程

离子的形式存在。

Quiz5 为什么丝氨酸蛋白酶一般存放在酸性的缓冲溶液中?

活性中心的 3 个氨基酸残基依次参与此阶段的反应,构成催化三元体结构:His 57 从 Ser195 移去一个质子,致使 Ser195 的亲核性大增,而 His 57 成为带正电荷的共轭酸。Asp102 通过氢键与 His 57 的咪唑基结合,其功能始终是定向 His 57,影响 His 的 pK_a。

(3) 碳四面体过渡态因氧阴离子在氧阴离子穴与 Ser195(—NH)及 Gly193(—NH)形成氢键而得以稳定。这些相互作用的净效应导致活化能的降低,可能对酶催化的贡献最大(图 10-15)。

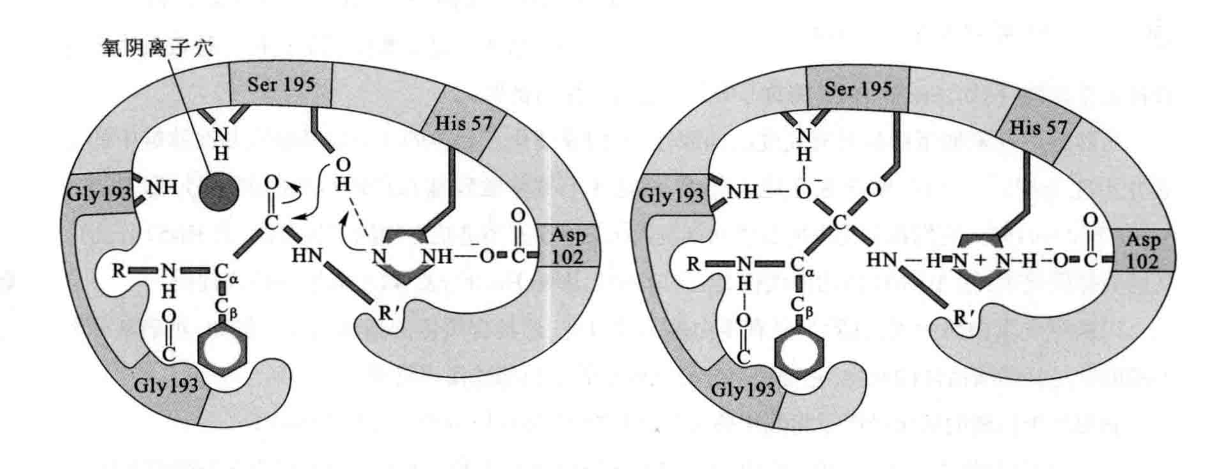


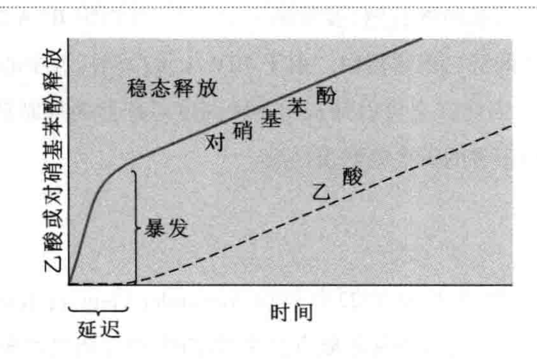
图 10-15 胰凝乳蛋白酶催化的四面体中间物的形成及其稳定

(4) 肽键断裂　原来肽键右侧的肽段作为离开基团，从 His57 咪唑环上得到一个质子后被释放，而原来肽键左侧的肽段则通过氢键和与 Ser195 形成的共价键仍然与酶结合。

(5) 水进入活性中心并质子化 His57，使之成为酸，被释放出来的 OH^- 亲核进攻留下来的多肽羰基碳，于是第二个四面体形的过渡态中间物形成了。最后，Ser195 从 His57 重新得到质子，His57 仍然与 Ser195 通过氢键相连，于是过渡态中间物发生瓦解，脂酰化酶中间物被水解，残留的多肽部分（第二个产物）被释放。

此阶段的反应为多肽水解的限速步骤，直接的证据来自使用对硝基苯酚乙酸酯为底物的酶动力学研究。在高浓度底物的存在下，产物的释放明显分为两个阶段：先是对硝基苯酚的暴发式（burst）释放，在经过一段时间延迟以后，才有产物乙酸缓慢的释放（图 10–16）。

Quiz6　丝氨酸蛋白酶共使用了哪几种催化机制？哪一种贡献最大？



E—OH + 对硝基苯酚乙酸酯（Ser 195）—快步骤→ E—O—C(=O)—CH$_3$（Ser 195）+ 对硝基苯酚负离子 + H^+；—慢步骤（H_2O，H^+）→ $^-$O—C(=O)—CH$_3$

图 10–16　以对硝基苯酚乙酸酯为底物的胰凝乳蛋白酶反应历程及动力学曲线

对硝基苯酚（相当于肽键右侧的氨基酸残基）的暴发式释放表明，肽键的右侧部分作为反应的第一个产物首先被游离出来，而乙酸（相当于肽键左侧的氨基酸残基）的延迟释放则意味着，在肽键断裂以后，它仍然与酶活性中心紧密地结合在一起，其释放速率决定了多肽水解的总速率。

2. 巯基蛋白酶

这一类蛋白酶也被称为半胱氨酸蛋白酶，它们广泛存在于自然界中。属于植物来源的有木瓜蛋白酶（papain）和菠萝蛋白酶（bromelain），其他的还有钙蛋白酶（calpain）、梭菌蛋白酶（clostripain）、小核糖核酸病毒蛋白酶（picornavirus proteinase）、链球菌蛋白酶（streptococcal proteinase）、白细胞介素 -1-β-转化酶（interleukin-1-β-converting enzyme）和一类参与细胞凋亡的胱天蛋白酶（cysteine aspartic acic specific protease，Caspase）。

巯基蛋白酶的共同特征是在催化过程中，需要一个 Cys 残基和一个 His 残基。其作用机制与丝氨酸蛋白酶有许多相似之处，而两者的主要差别在于巯基比羟基更容易发生去质子化。当巯基蛋白酶活性中心的巯基将质子交给相邻的 His 残基以后，带负电荷的 S 作为强亲核基团，进攻肽键的羰基碳，形成以硫酯键连接的脂酰化酶共价中间物。余下的反应与丝氨酸蛋白酶催化的反应相似。

Quiz7　某些巯基蛋白酶上起亲核催化的 Cys 被 Ser 取代以后仍然保持着活性，某些则不行，为什么？

（二）天冬氨酸蛋白酶和金属蛋白酶

这两类蛋白酶催化的反应比较直接，只有一次亲核进攻，且亲核试剂都是水分子，因此没有共价

e10-1 天冬氨酸蛋白酶和金属蛋白酶

催化。其中天冬氨酸蛋白酶的催化涉及两个重要的 Asp 残基，它们协调一致，交替充当广义的酸碱催化剂，而金属蛋白酶的活性绝对需要金属离子。金属离子的作用可能是激活水分子，或者进行路易斯酸催化。

对于大多数天冬氨酸蛋白酶而言，都具有由两叶（每一叶相当于一个结构域）组成（N 端和 C 端各成一叶）的三级结构。两叶差不多呈两重对称，少数为同源二聚体，例如胃蛋白酶、组织蛋白酶 D（cathepsin D）、凝乳酶（chymosin）、肾素和艾滋病病毒 -1（HIV-1）蛋白酶。这些天冬氨酸蛋白酶在中性或偏酸的条件下活性最高，与丝氨酸蛋白酶一样，在功能上要么属于消化酶，如胃蛋白酶和凝乳酶，要么属于调节酶，如肾素参与调节血管紧张素的活性。

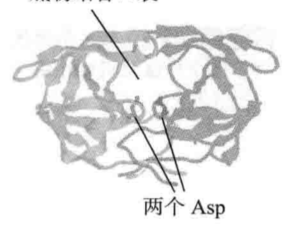


图 10-17 HIV 蛋白酶的二聚体三维结构

HIV-1 蛋白酶是一种比较新颖的天冬氨酸蛋白酶，它在 HIV-1 生活史中负责切开其基因组编码的多聚蛋白质产物，促进新病毒颗粒的成熟（参看第三十八章“基因组 RNA 的复制”）。与其他天冬氨酸蛋白酶不同的是，它是一个同源二聚体蛋白。由于 HIV-1 蛋白酶活性中心高度对称性（与其他天冬氨酸蛋白酶相比）（图 10-17），因此完全可以筛选出一种高度对称性的抑制剂——仅仅抑制 HIV-1 蛋白酶活性，而不会抑制人体内正常的天冬氨酸蛋白酶。

二、溶菌酶

溶菌酶广泛存在于动、植物中，早在 1922 年就由 Alexander Fleming 发现。人溶菌酶主要分布在眼泪、血液和唾液的白细胞之中。溶菌酶并不能直接裂解细菌，而是通过水解细菌细胞壁上的肽聚糖，导致细菌不能抵抗渗透压的变化而发生裂解（图 10-18）。肽聚糖是由短肽和多糖形成的共价复合物（参看第十五章“糖类”），其中的多糖链是由 *N*- 乙酰胞壁酸（NAM）和 *N*- 乙酰葡糖胺（NAG）通过 β-1，4 糖苷键相连。

图 10-18 溶菌酶催化的水解反应

鸡蛋的蛋清是溶菌酶很好的来源，其中的溶菌酶含有 129 个氨基酸残基和 4 个二硫键。溶菌酶是第一个使用 X 射线得到三维结构的酶。研究表明，其活性中心是一个长而深的裂缝，正好可以容纳 6 个单糖单位。如此多的单糖单位结合在活性中心，意味着底物与活性中心之间形成了大量的氢键。

天然的底物在活性中心上不稳定，因此并不适合用来研究活性中心的结构。而使用天然底物的类似物——不同长度的 $(NAG)_n$（图 10-19）却能解决许多问题。

图 10-19 $(NAG)_3$ 的化学结构

按照这样的思路，给溶菌酶提供不同长度的寡聚 NAG（各个单糖残基从非还原端到还原端的次序，按 A、B、C、D、E、F 等的次序编号），水解的

结果见表 10–2：

► 表 10–2 溶菌酶对不同底物水解的相对速率

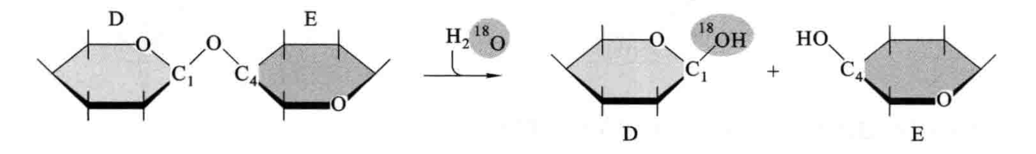
$(NAG)_n$	水解的相对速率	$(NAG)_n$	水解的相对速率
$(NAG)_2$	0	$(NAG)_5$	4 000
$(NAG)_3$	1	$(NAG)_6$	30 000
$(NAG)_4$	8	$(NAG)_8$	30 000

表中的数据显示，$(NAG)_2$ 和 $(NAG)_3$ 非常稳定，这意味着 A–B 和 B–C 之间的糖苷键不大可能水解。由于酶分子上与 C 结合的位点上不适合结合 NAM，因此，细菌细胞壁上合适的底物序列一定是：

NAG—NAM—NAG—NAM—NAG—NAM

A　B　C　D　E　F

如果要让 NAG 更好地与酶活性中心上的 D 位点结合，就需要糖环的扭曲。这里涉及酶诱导的底物形变。因为溶菌酶只能特异性地切开 NAM—NAG 之间的键，所以 C—D 和 E—F 之间的键不能被水解。通过排除法，能水解的位点只能是 D—E 之间的糖苷键了。那么，D—E 之间的哪一个碳氧键被切开了呢？使用 ^{18}O 标记的水可以巧妙地解决这个问题。由于碳氧键断裂以后，^{18}O 只出现在 C1 位置上，因此可以推断，被切开的碳氧键是 C1–O。



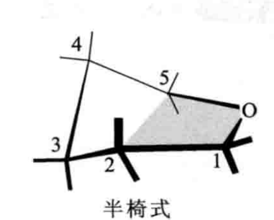
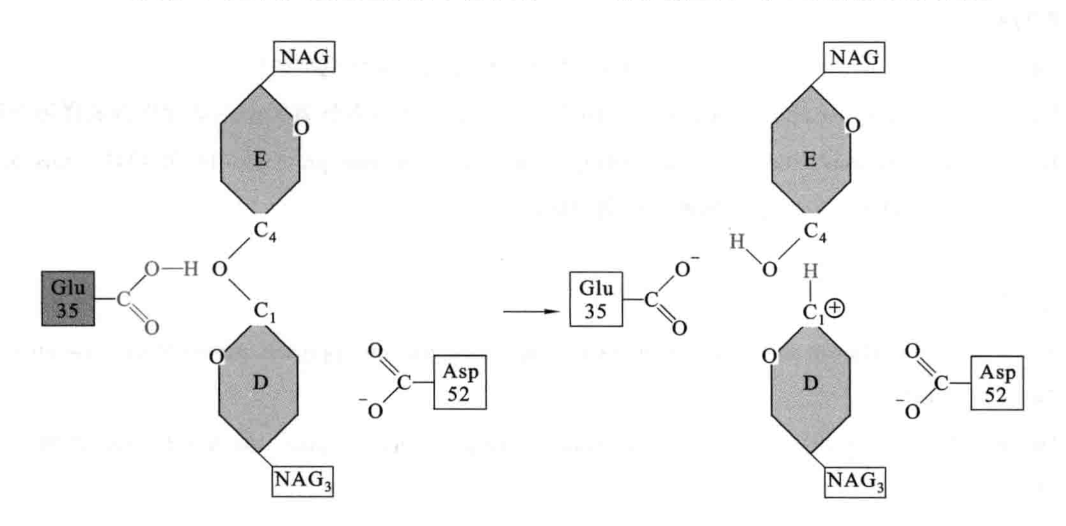


图 10–20 溶菌酶催化过程中 D 位点糖环的构象变化

综合其他方面的研究成果，现在认为溶菌酶的作用机制可分为以下几步：

1. Glu35 作为广义酸将其羧基上的质子提供给 D—E 糖苷键上的 O，导致糖苷键的断裂。



2. D 位点上的 C1 带上正电荷，成为一个不稳定的碳正离子中间物。这种不稳定的中间物因为 D 位点糖环的扭曲（从椅式变成半椅式）（图 10–20）以及 Asp52 所带的负电荷的中和作用而得以稳定。

$(NAG)_4$ 内酯类似物（lactone analogue）的 D 环也是半椅式构象（图 10–21），与过渡态十分相似，因此它是溶菌酶的强抑制剂。

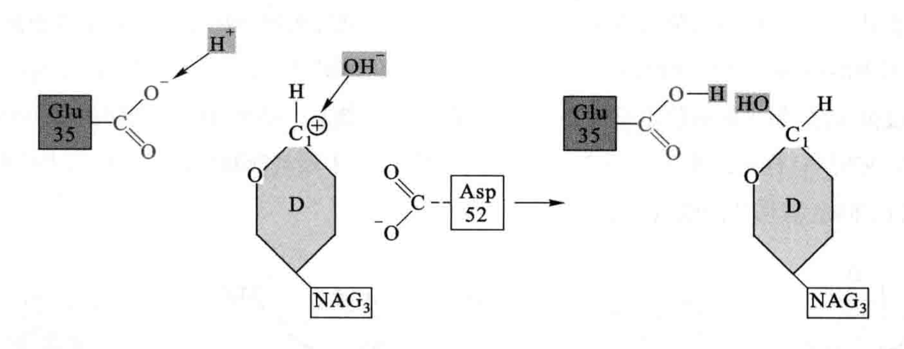
(NAG)$_4$碳正离子衍生物
(过渡态中间物)

(NAG)$_4$内酯类似物
(过渡态类似物)

图 10-21　溶菌酶的过渡态中间物和过渡态类似物

3. E-F 二糖单位通过扩散离开酶活性中心。

4. 碳正离子中间物与溶剂中的 -OH 反应，产生（NAG）$_4$。



5.（NAG）$_4$ 扩散离开酶活性中心，Glu35 重新质子化。

Quiz8 溶菌酶共使用哪几种催化机制？

e10-2 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Enzyme_catalysis（维基百科有关酶催化机制的内容）
2. http://www.cryst.bbk.ac.uk/PPS2/course/section12/enzymes.html（内有丝氨酸蛋白催化机制的详细介绍）
3. http://www.nature.com/scitable/topicpage/enzyme-catalysis-the-serine-proteases-14398894（*Nature* 杂志图书馆提供的有关丝氨酸蛋白酶催化机制的最新内容）

参考文献：

1. Puente X S, et al. Human and mouse proteases: a comparative genomic approach. *Nature Reviews Genetics*, 4: 544–558, 2003.
2. De Prada P, et al. Application of artificial enzymes to the problem of cocaine. *Ann N Y Acad Sci*, 909: 159–169, 2000.
3. Kraut J. How do enzymes work? *Science*, 242: 533–540, 1988.
4. Carter P, Wells J A. Dissecting the catalytic triad of a serine protease. *Nature*, 332: 564–568, 1988.

第十一章 核酶的结构与功能

核酶就是 RNA 催化剂。它的发现意义重大,不但纠正了人们认为酶都是蛋白质的传统观念,而且对探索生命的起源很有启发意义。在漫长的生命进化过程中,地球上很可能曾出现过一个奇特的由 RNA 独领风骚的"RNA 世界"。那时,既无 DNA,又无蛋白质,但有 RNA。在 RNA 世界里,RNA 不仅充当遗传物质,还行使催化功能。然而,随着生命的不断进化,原始 RNA 的两项功能分别"让位"给了 DNA 和蛋白质,那些至今仍然保留催化活性的 RNA 似乎是这种进化历程留下的"活化石"。目前有很多证据支持这种学说的正确性,事实上,许多重要的证据就隐藏在现代的活细胞内部。

既然核酶也属于酶,就应该具有酶的一切性质。但作为一种最古老的生物催化剂,核酶催化反应的种类和效率与基于蛋白质的酶相比,还存在一些不足。正因为如此,现代的细胞主要使用蛋白质作为催化剂,但少数核酶仍然被使用,是因为它们催化的反应已能完全满足细胞功能的需要,在生命的进化过程中,缺乏被能行使相同功能的蛋白质取代的选择性压力。

本章将主要介绍核酶的主要类型、催化机制与应用以及支持"RNA 世界"学说的主要证据。

第一节 核酶的种类

到目前为止,已发现近十种天然的核酶。这些天然的核酶在细胞里催化不同的反应。根据分子的大小和催化机制,核酶可以分为小核酶和大核酶,其中小核酶的大小 40~154 nt,大核酶有几百个到几千个 nt。

e11-1 形形色色的天然核酶

小核酶一般来源于某些动、植物病毒的卫星 RNA,主要包括锤头状(hammerhead)核酶、发夹状(hairpin)核酶、D 型肝炎病毒(hepatitis delta virus,HDV)RNA、Varkud 卫星(Varkud satellite,VS)核酶和 GlmS 核开关。这些核酶除了 GlmS 核开关以外,都能催化近乎相同的剪切反应,在宿主细胞内都是通过互补的 RNA 中间物进行滚环复制(图 11-1),但在二级结构和三级结构上差别很大。滚环复制产生的是多个拷贝串联在一起的多聚体,需要通过自我剪切的方式将单体释放出来。

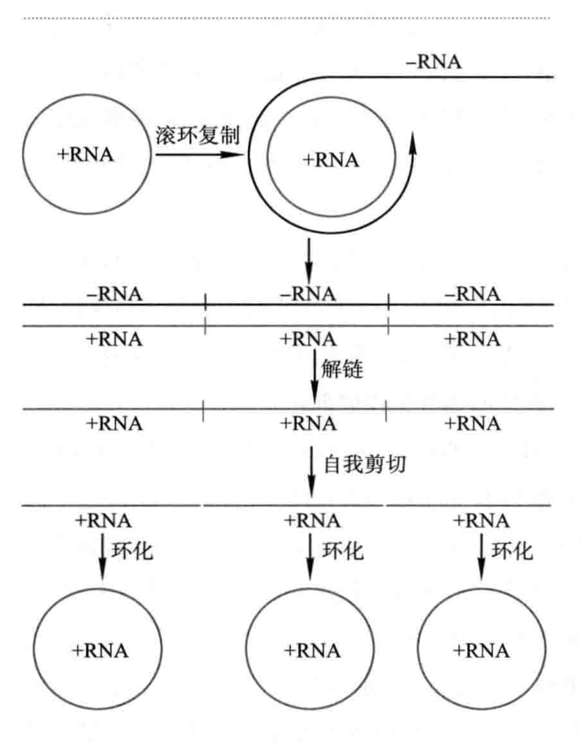


图 11-1 RNA 的滚环复制以及子代 RNA 的成熟

大核酶包括Ⅰ组和Ⅱ组自我剪接的内含子(self-splicing introns)、催化真核细胞核 mRNA 前体剪接的剪接体(spliceosome)、催化 tRNA 前体在 5′端后加工的核糖核酸酶 P 和催化蛋白质生物合成的核糖体。

Quiz1 哪些天然核酶存在于所有生物的体内?哪些存在于所有的真核生物体内?

除了以上这些天然的核酶以外,人们使用指数富集式配体系统进化(SELEX)技术,已从人工构建的序列库中筛选到自然界并不存在的核酶。在体外催化反应的类型上,这些人工核酶大大超过了细胞内现存的天然核酶,至少包括:催化 RNA 复制的 RNA 复制酶;催化 RNA 转录的 RNA 聚合

框 11-1　生化热点——DNA 也能催化吗?

在各种天然核酶和人工核酶被发现以后,科学家们一直在思考,是否有 DNA 也能行使催化? 由于细胞内 DNA 绝大多数都是以双螺旋结构存在的,这种千篇一律的结构限制使其不大可能与底物结合并催化反应。然而,人工合成的单链 DNA 缺少了双链 DNA 在结构上的限制,可以像单链 RNA 一样折叠成丰富多彩的三维结构,会不会少数也有催化功能呢?

1994 年,Breaker R R 等人使用 SELEX 技术进行体外筛选(in vitro selection)(参看第四十四章"重组 DNA 技术"第八节内容),从随机建立的单链 DNA 库中,率先得到了一种具有催化性质的人工 DNA,该 DNA 能在 Pb^{2+} 存在的情况下,催化 RNA 的剪切。在随后的十几年中,又有多种具有催化活性的人工 DNA 被筛选出来,催化的反应包括 DNA 连接、RNA 连接、DNA 磷酸化、DNA 腺苷酸化、DNA 剪切、DNA 的水解、卟啉金属化、嘧啶二聚体的修复、核肽键的形成(nucleopeptide bond)以及碳—碳键的形成。这些具有催化活性的 DNA 也可被称为脱氧核酶(deoxyribozyme)。在这些由 DNA 催化的反应中,底物多数是 DNA,原因是 DNA 本身就有与其他核酸通过互补碱基对结合的内在性质。

然而,迄今为止,还没有在任何生物、任何细胞中发现有催化活性的天然 DNA,这与核酶是不一样的。

酶;催化 ATP 的 γ- 磷酸基团转移到 RNA 5′- 或 2′-OH 的多核苷酸激酶;催化氨酰基从 Phe-AMP 转移到 RNA 3′端的氨酰 -RNA 合成酶;催化 C—N 键和 C—S 键形成的酶;催化酰胺键形成的酶;催化碱基和戊糖形成核苷酸的酶;催化尿苷酸甲基化形成胸苷酸的合酶;催化脂酰 -CoA 合成的合成酶等。这些核酶很可能在生命的进化过程中出现过,属于从 RNA 世界过渡到现代的生命世界中失落的环节。

第二节　核酶的催化机制

核酶与起催化作用的蛋白质具有许多相同或相似的特征,包括只改变反应速率,不改变平衡常数;具有很高的底物专一性;对温度和 pH 等因素的变化敏感;具有较高的催化效率,例如,HDV 核酶的周转数为 10^2~10^4 s^{-1},这与核糖核酸酶 A 不相上下(~10^3 s^{-1});也可以使用辅助因子来协助催化,例如,绝大多数核酶需要金属离子为辅助因子,Ⅰ组内含子使用鸟苷或鸟苷酸为辅助因子;酶活性既可以受到激活,也可以受到抑制,例如,GlmS 核开关能受到 6- 磷酸葡糖胺的激活,也能受到 6- 磷酸葡糖的竞争性抑制;酶的活性也依赖于其特有的三维结构,特别是活性中心的结构,并且也是通过稳定反应的过渡态、降低反应的活化能来提高反应速率的。

Quiz2 你认为核酶的动力学曲线是双曲线还是 S 型? 为什么?

然而,与基于蛋白质的酶相比,天然核酶能催化的反应较为单调。几乎所有已知的天然核酶催化的都是与磷酸二酯键有关的转酯反应,功能一般为 RNA 的后加工。而且,大多数核酶属于自催化,即催化自身分子内的反应。这与酶的确切含义有别,原因是真正的酶在催化反应结束以后,自身的结构不应该有变化。核酶催化的反应有限,这是因为其活性中心能够提供的催化基团很有限。蛋白质共有 10 种蛋白质氨基酸的侧链具有不同的 pK_a,特别是 His,其侧链的 pK_a 接近生理 pH。然而,RNA 只有 4 种碱基组成,且大都埋在具有高度的负电荷的磷酸核糖骨架上。碱基的 pK_a 要么太低(A 和 C 分别是 3.5 和 4.2),要么太高(G 和 U 分别是 9.8 和 10.5)。在核酶的活性中心,碱基一般很难有效地作为广义的酸碱催化剂去参与催化。正因为如此,几乎所有天然的核酶(可能的例外是发夹状核酶)都绝对需要二价的金属离子 Mg^{2+} 或 Mn^{2+}。这两种金属离子所起的作用可能是结构上的,也可能直接参与催化。金属离子起结构性作用的典型例子是四膜虫的Ⅰ组内含子核酶,它的一个亚结构域的折叠依赖于 3 个 Mg^{2+} 与 2 个磷酸基团的结合。这种结合使得 RNA 能够形成里朝外或内翻的结构,即磷酸基团伸向内部,碱基朝外指向溶剂;金属离子参与催化的典型例子也是四膜虫的Ⅰ组内含子,其活性中

心绝对需要 2 个 Mg^{2+} 参与催化。

金属离子要是参与催化，那所起的作用要么是激活亲核基团，要么是稳定转酯反应内的离去基团。例如，Ⅰ组内含子活性中心的 2 个 Mg^{2+}，一个激活亲核基团，一个稳定离去基团。然而，核酶并不限于金属催化，某些核酶也可能使用广义的酸碱催化。对于这些核酶而言，其活性中心的碱基或核糖羟基，甚至主链上的磷酸基团都可作为质子供体或受体参与催化。还有的核酶使用其他辅助因子来协助催化，以弥补活性中心催化基团的不足。下面就分别简要介绍小核酶和大核酶的催化机制。

一、小核酶的催化机制

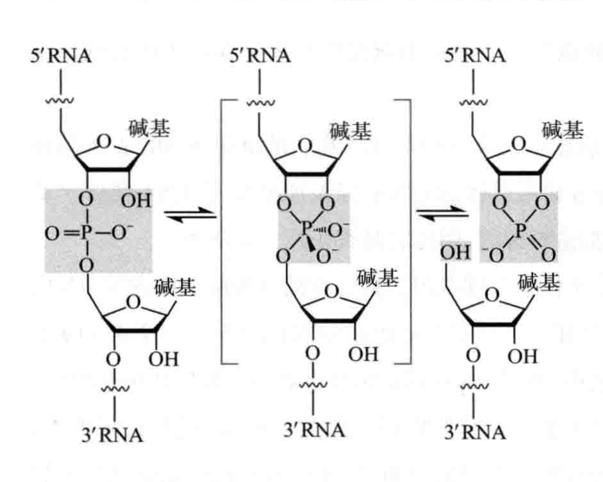


图 11-2 小核酶所催化的反应的基本步骤

这一类核酶催化的都是位点特异性剪切 / 连接（cleavage/ligation）反应，催化机制都涉及到一个被激活的亲核基团对一个磷酸二酯键的进攻，形成五价磷过渡态或半衰期极短的中间物，然后是一个离去的氧原子。反应的结果是磷酸基团的立体化学发生变化。这些核酶在催化机制上的差别主要是亲核基团和离去基团的不同。如图 11-2 所示，4 种小核酶都使用内部紧靠剪切点的一个核苷酸上的 2′-OH 为亲核基团，去取代剪切点上的磷酸二酯键的 5′-O，反应产物为带有 2′，3′-磷酸二酯键的 RNA 和一个游离的 5′-OH。这样的反应与 RNA 的碱水解很相似，但核酶催化的是特定位置的单个磷酸二酯键的水解。除了发夹状核酶以外，其他都需要金属离子协助催化。金属离子的结合位点是高度特异性的，不可以用其他带正电荷的离子代替。

以锤头核酶为例，这一类核酶属于天然核酶中最小的，只有约 40 nt，因二级结构形如锤头（图 11-2）而得名。所有锤头核酶都含有三个碱基配对的茎（茎Ⅰ、茎Ⅱ和茎Ⅲ）和一个高度保守的为切割所必需的 CUGA 序列。三个茎的序列可以变化，但必须维持 Watson-Crick 碱基对。中央核心区序列在各种锤头核酶内是高度保守的。

中央核心区包括两个结构域：结构域Ⅰ由 CUGA 序列和尿苷转角（the uridine turn）组成，紧靠茎Ⅰ；结构域Ⅱ由非 Watson-Crick 碱基对组成，将茎Ⅱ和茎Ⅲ连接起来。

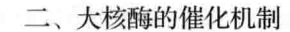
剪切反应开始于催化位点上的胞苷酸 2′-OH 上的氧对邻近核苷酸 3′-磷酸基团的亲核进攻，Mg^{2+} 在其中也起作用。亲核进攻最终导致 3′，5′-磷酸二酯键的断裂，同时产生 2′，3′-环形核苷酸残基（图 11-3）。此外，锤头状核酶活性中心有两个不变的 G，其中一个 G（G12）的 N1 靠近 2′-OH，另一个 G（G8）

图 11-3 锤头核酶的二级和三级结构以及可能的催化机制

的 2′-OH 靠近 5′-O 离去基团。这两个 G 也可能参与催化，其中第一个 G 激活亲核基团，第二个 G 稳定离去基团。

除了锤头核酶以外，发夹状核酶、HDV 核酶和 GlmS 核开关核酶也属于小核酶。其中，发夹状核酶来自一些植物病毒的卫星 RNA，如马铃薯环斑病毒（tobacco ringspot virus）卫星 RNA、菊苣根黄斑病毒（chicory yellow mottle virus）卫星 RNA 和南芥菜花叶病毒（arabis mosaic virus）卫星 RNA；HDV 为人乙型肝炎病毒（HBV）的卫星病毒；Glm 核开关来自枯草杆菌，位于编码 6- 磷酸葡糖胺（glucosamine-6-phosphate，GlcN6P）合成酶的 mRNA 在 5′端的非翻译区。

e11-2 发夹状核酶、HDV 核酶和 GlmS 核开关核酶



二、大核酶的催化机制

大核酶催化的反应有剪切反应、剪接反应和转肽反应。其中最典型的代表是存在于所有细胞中的核糖核酸酶 P。

核糖核酸酶 P 催化的是 tRNA 前体在 5′端的后加工反应（图 11-4）。在它的催化下，tRNA 的前体在 5′端某一特定的位置被引入一个切口，产生带有 5′磷酸基团的 tRNA，以及原来位于 tRNA 前体 5′端的一段前导序列。鉴于核糖核酸酶 P 使用的底物都是外源的，因此它是真正意义上的酶。

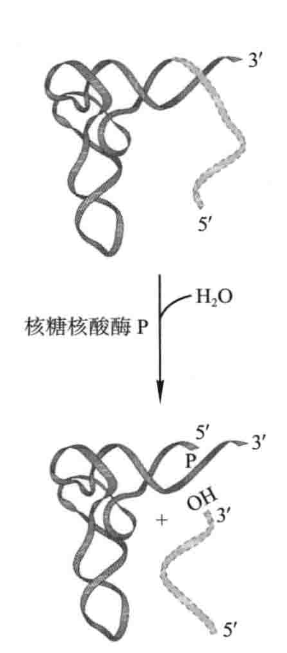


图 11-4 核糖核酸酶 P 催化的剪切反应

与其他核酶不同的是，核糖核酸酶 P 使用水分子作为亲核基团，并且，核糖核酸酶 P 既含有 RNA，又含有蛋白质。例如，大肠杆菌的核糖核酸酶 P 含有一个由 377 nt 组成的 M1 RNA 和一个由 119 个氨基酸残基组成的 C5 蛋白。Sidney Altman 的研究表明，M1 RNA 在高盐浓度下能独立地完成催化反应，但有 C5 蛋白可使反应速率提高 2~3 个数量级。C5 蛋白所起的作用是作为一种碱性蛋白，降低 M1 RNA 与 tRNA 底物之间的静电排斥力，即提高核酶与底物之间的亲和力，还可能通过增强金属离子与酶活性中心的亲和性而协助催化，以及促进产物的释放。

2010 年 9 月美国西北大学的 Nicholas J. Reiter 等人，分别得到了海栖热袍菌（*Thermotoga maritime*）的核糖核酸酶 P 全酶在两种情况下的晶体结构（图 11-5）：一种情况仅仅是与成熟的 $tRNA^{Phe}$ 形成的复合物，另外一种情况是多了 tRNA 前体在 5′端被切下的前导序列。从晶体结构中可以看出，核糖核酸酶 P 的蛋白质部分与底物，即 tRNA 前体长达 59 nt 的前导序列结合，而 RNA 部分（P RNA）在多个位点与 $tRNA^{Phe}$ 发生作用，包括 P RNA 的 S 结构域与 tRNA 的 TΨC 环和 D 环之间的碱基堆积、在 tRNA 受体茎内的 A- 小沟作用，以及与 tRNA 的 3′端几个碱基形成的标准碱基对。整个 P RNA 共有两个结构域，一个是特异性结构域（specificity domain，S 结构域），另一个是催化结构域（catalytic domain，C 结构域），两个结构域通过一个共轴螺旋相连，其中涉及多个远程作用，包括所有生物 P RNA 共同的 5 个保守的区域（CR-I~CR-V）。这些保守的区域形成两个功能区，一个区参与识别底物，另一个区形成活性中心的结构框架。

Quiz3 当初在 Altman 宣布核糖核酸酶 P 是核酶以后，很多怀疑论者认为，他们分离到的 RNA 仍然有痕量的蛋白质存在，正是残留的蛋白质在起催化作用。对此，试设计一个实验来支持 Altman 的结论。

核糖核酸 P 的催化机制是依赖于 2 个 Mg^{2+} 的双金属催化（图 11-6），1 个 Mg^{2+}（Mg_1^{2+}）激活充当亲核试剂的羟基，使这个羟基进行亲核取代，形成一个新的键，取代在 3′端切开的磷酸基团上的氧。另 1 个 Mg^{2+}（Mg_2^{2+}）与水分子结合，促使质子转移给在 3′端切开的磷酸基团上的氧，从而促进产物的释放。

e11-3 Ⅰ组和Ⅱ组内含子

除了核糖核酸酶 P 以外，Ⅰ组和Ⅱ组内含子、剪接体（spliceosome）和核糖体也属于大核酶。其中，Ⅰ组和Ⅱ组内含子催化自我剪接反应，剪接体催化真核细胞核 mRNA 前体的剪接反应（参看第三十七章“转录后加工”），核糖体催化翻译中肽键的形成，即转肽反应（参看第三十九章“蛋白质的生物合成及其在细胞内的降解”）。

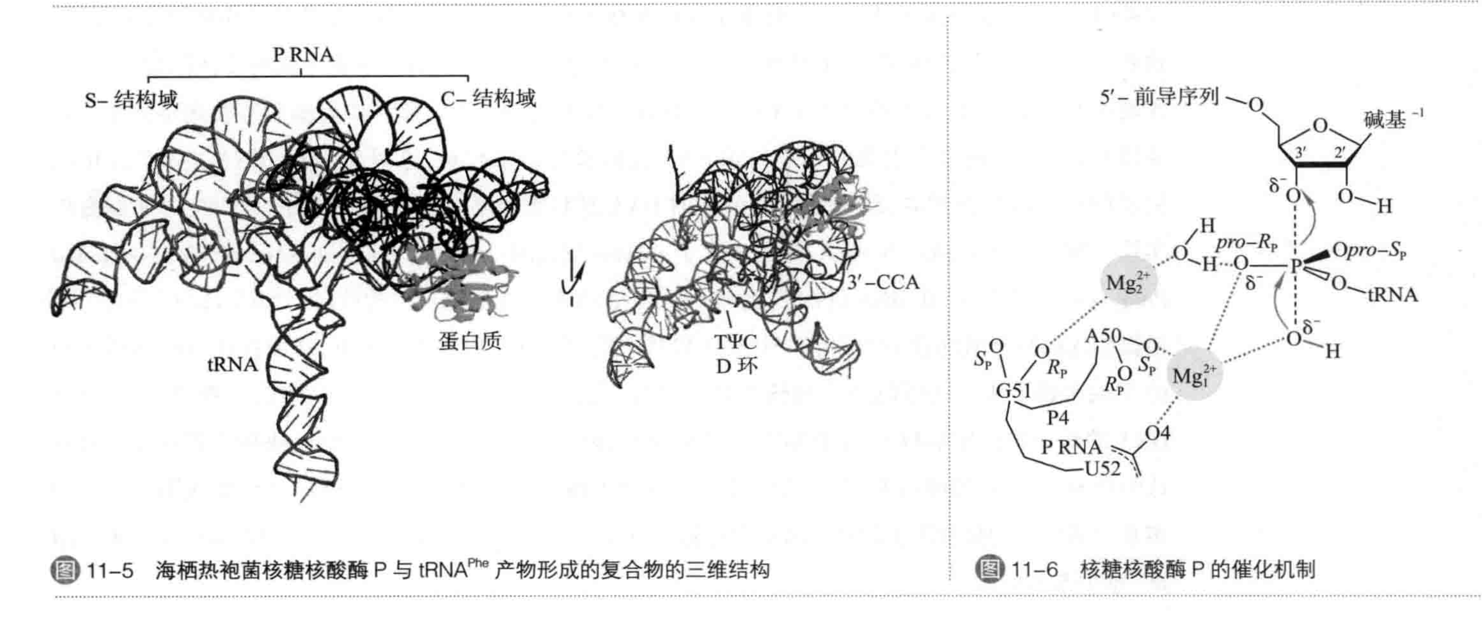


图 11-5 海栖热袍菌核糖核酸酶 P 与 $tRNA^{Phe}$ 产物形成的复合物的三维结构

图 11-6 核糖核酸酶 P 的催化机制

第三节 核酶发现的意义及其应用

有关地球上生命起源的问题一直是众多科学之谜中的一个，而在生命进化过程中，究竟是先有 DNA，还是先有蛋白质？这个类似“先有鸡，还是先有蛋”的问题，曾长期在生命科学界争论不休。直至核酶被发现以后，再加上细胞内多种具有不同结构和功能的 RNA 被相继发现，人们终于意识到在生命的进化过程中，RNA 占据了十分特殊的位置。在地球上曾经出现过一个奇特的“RNA 世界”，在那个世界里，RNA 既充当遗传物质，又行使酶的功能，催化自身的复制和提供原始生命所需要的其他催化活性。显然，它的产生先于蛋白质和 DNA（图 11-7）。然而，随着生命的进化，生物体内的代谢越来越复杂，对生物催化剂的要求就越来越高，RNA 作为生物催化剂固有的不足已不能满足需要。这种选择性压力要求出现更快、更强、更好的催化剂，于是基于蛋白质的生物催化剂出现了，并逐步取代了那些已不能满足生命更高要求的核酶，而少数至今保留下来的核酶显然仍然能够完全胜任现代细胞所要求的工作，故还没有被取代的必要。

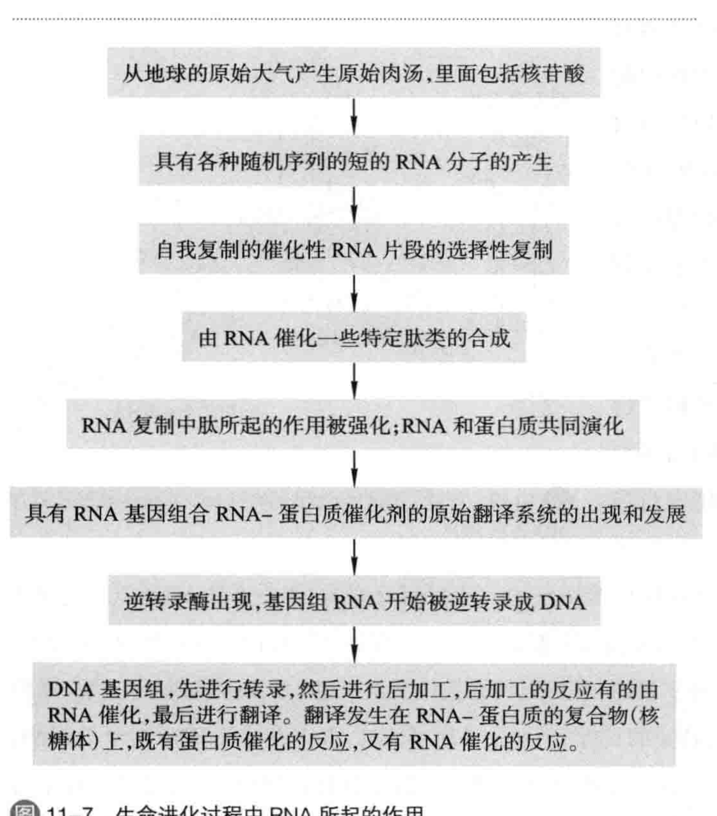


图 11-7 生命进化过程中 RNA 所起的作用

目前已积累了大量支持 RNA 世界学说的证据，主要包括：① RNA 也可以充当遗传物质，很多病毒（如 HIV 和流感病毒）以 RNA 为遗传物质。② RNA 可以充当酶，而且可以催化多种不同的反应。现代细胞里，还残留的天然核酶能够催化蛋白质的合成（核糖体）、催化转录后加工（核糖核酸酶 P、Ⅰ组和Ⅱ组内含子、剪接体）、参与 RNA 的复制（垂体状核酶、发夹状核酶和 HDV 核酶）、参

Quiz4 如果你被派往其他星球，去寻找有没有生命，你应该根据哪一种分子的有无来判断外星生命是否存在？

与基因表达的调控(GlmS 核开关)。而通过 SELEX 技术制备筛选出来的核酶催化的反应种类就更多了，这些人工核酶也许曾经在 RNA 世界中出现过，只不过因为后来跟不上生物进化的步伐而被淘汰了。③现代的细胞里，有各种不同的 RNA，它们具有多项不同的功能。除了作为酶和某些病毒的遗传物质以外，还有其他多个重要的功能。④ RNA 也能参与转录校对，这对于当初“RNA 世界”的 RNA 复制的忠实性极为重要，就像现代的细胞的 DNA 复制需要通过 DNA 聚合酶的自我校对来提高忠实性一样。根据 Zenkin N. 等人在 2006 年于 *Science* 发表的一篇论文 *Transcript-Assisted Transcriptional Proofreading* 的结果，在 DNA 进行转录的时候，RNA 事实上在催化自己的校对。如图 11-9 所示，一个错误参入的核苷酸能够折回到前一个核苷酸的位置，在 Mg^{2+} 和水的存在下，从而催化自己从转录物的 3′端水解下来。⑤细胞在合成核苷酸的时候，先合成 RNA 的组成单位核糖核苷酸，然后再合成 DNA 的组成单位脱氧核苷酸，即脱氧核苷酸是由核糖核苷酸还原而来。此外，细胞是先合成一般被认为是 RNA 特有的碱基 U，然后再合成 DNA 特有的碱基 T(参看第三十二章“核苷酸代谢”)。⑥许多基于蛋白质的酶所需要的辅助因子含有腺苷或腺苷酸，如 FAD、NAD^+、$NADP^+$ 和 CoA(参看第十四章“维生素与辅酶”)。

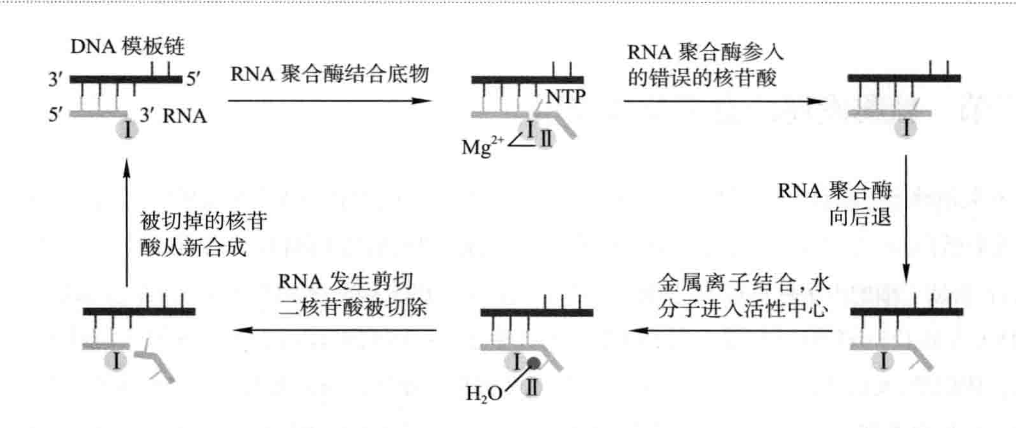


图 11-8 RNA 协助的转录校对(Campbell 等，2009)

核酶的发现对理解生命的起源和进化具有十分重要的理论意义。此外，人们在对核酶的作用机制有了深入理解的基础上，可以将其改造成一种能在体内定向切割特定目标 RNA 分子的工具，并应用于多项研究，特别是医学研究之中。虽然大多数天然的核酶催化的是自切割反应，但是它们的活性都依赖于各自独特的一级、二级结构和三级结构，特别是二级结构和三级结构。同时，它们作用的机理主要依赖于特异性序列通过互补碱基对来识别并结合特异性目标序列，因此完全可以在得到一个特定目标 RNA 的序列之后，按照一种核酶对一、二、三级结构的基本要求(如锤头状核酶和发夹状核酶剪切点的一致序列分别为 GUC*N 和 YN*GUC，这里 N 表示可以是四种核苷酸中的任何一种，Y 为嘧啶核苷酸，* 为切点)，在体外设计出另外一种 RNA 分子。一旦这种 RNA 分子与特定目标 RNA 分子相遇，就可以特异性结合在一起，并自组装成特定核酶活性所需要的二级结构和三级结构。于是，特定的目标 RNA 分子充当底物，人工设计的 RNA 分子充当酶，在它们结合以后，目标 RNA 分子就可以在特定的位点被切开。按照这样的思路，人们已经设计了各种各样的核酶，主要有锤头状核酶、发夹状核酶和 HDV 核酶，它们作用的对象都是高度特异性的，可

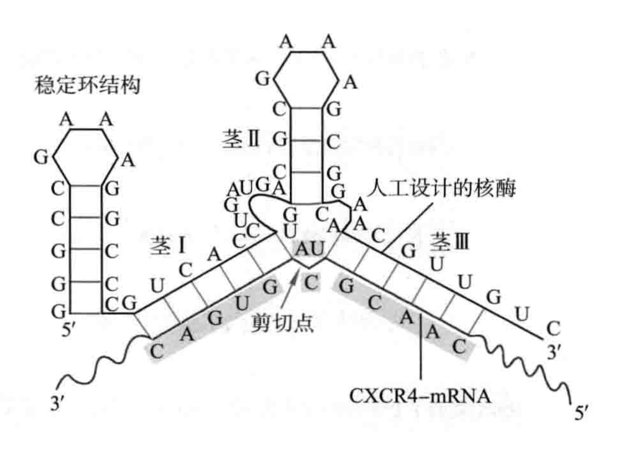


图 11-9 计算机辅助设计出来的抗 HIV 感染的核酶及其作用点目标 RNA

以是各种 RNA 病毒基因组上特定的序列，也可以是细胞内特定的 mRNA。利用这些核酶，人们可以对付各种病原体，进行基因治疗研究特定基因的功能。当由一个特定基因编码的 RNA 被使用的核酶切割以后，必然影响到该基因的功能，并可能给生物的表型带来某种变化，借此可以推断出原来基因的功能，这与基因敲除和基因敲减技术(参看第四十四章“重组 DNA 技术”)具有异曲同工之妙。例如，HIV 的引物结合位点(primer binding site，PBS)是一个 18nt 的高度保守序列，它和宿主细胞的 $tRNA^{Lys}$ 的 3′端互补。Amer M. Kechli 等人针对这段序列设计了锤头状核酶，成功地抑制了 HIV-1 在体内的复

框 11-2　生化新见解——“病毒创造了 DNA”？

按照“RNA 世界”的学说，生命在进化的过程中，是先有 RNA，后来才有 DNA 和蛋白质。然而，DNA 是如何产生的，又是如何取代 RNA 而成为主要的遗传物质的呢？

“RNA 世界”学说只是认为，DNA 比 RNA 稳定，在生命进化很早的阶段，取代了 RNA，并成为细胞内的遗传物质，因此，现在地球上三大生命系统——细菌、古菌和真核生物最后的共同祖先是含有 DNA 的细胞。但这种含有 DNA 的远古细胞是如何得到 DNA 的呢？对此，法国巴斯德研究所的 Patrick Forterre 提出了一个十分新颖的假说(图 11-10)。Forterre 认为，单凭遗传物质稳定性小小改进的这点好处，还不足于让生命系统选择放弃 RNA 改用 DNA 作为遗传物质，而是远古细胞里的病毒“发明”了 DNA 作为一种修饰的机制，以防止它们的基因组受到宿主细胞设计的降解系统的破坏。当今的病毒有 RNA 病毒和 DNA 病毒，而 DNA 病毒中的 DNA 有时含有 U 或羟甲基 C 代替 T。而且，现代的三大生命系统之中都存在有专门破坏外源 DNA 或 RNA 的系统。Forterre 认为，起源于 RNA 世界的原始细胞含有 RNA 基因组，还有 RNA 病毒。在进化中，具有 DNA 基因组的病毒得到选择，是因为 DNA 不会被当时细胞内的核酸酶降解。这一切发生在三大生命系统出现之前。共有三类非毒性的 DNA 病毒(Forterre 将其称为创立者(founder)病毒)感染三大生命系统的祖先，它们在宿主细胞中作为 DNA 质粒进行复制，其中有两种亲缘关系更近，分别感染古菌和真核生物的祖先，而另外一种感染的是细菌的祖先。慢慢地宿主细胞将它们的遗传物质转化成稳定性更高的 DNA，那时逆转录酶已经出现，在将 RNA 变成 DNA 过程中起十分关键的作用，就像它现在在逆转录病毒中所起的作用一样。由于 DNA 比 RNA 稳定，那些没有被 DNA 病毒感染的使用 RNA 基因组的细胞不会成为 DNA 细胞，它们最终便灭绝了。

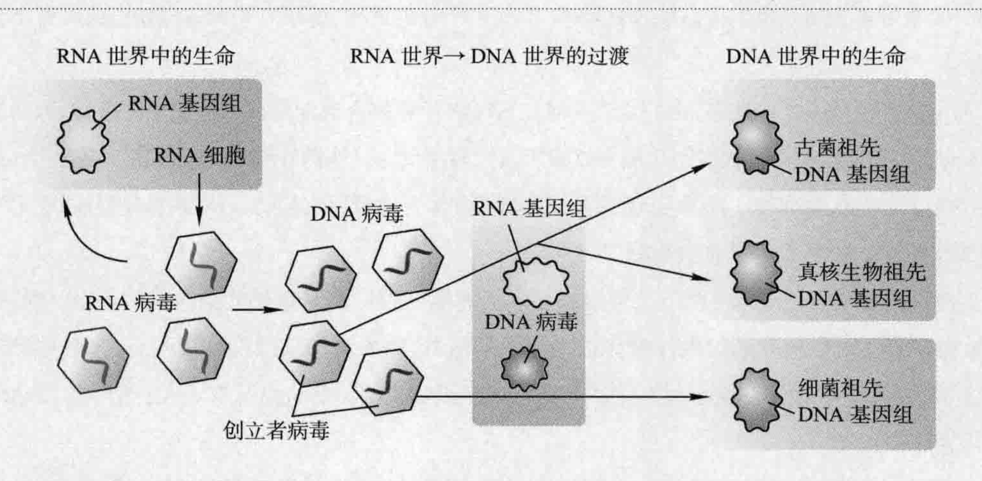


图 11-10　病毒创造 DNA 学说的图解(Madigan 等，2010)

Forterre 假说解释了细胞内 DNA 的起源，提供了一种 DNA 逐步取代 RNA 基因组的机制。重要的是，假说也解释了相比于转录和翻译，不同生命系统中的 DNA 复制、重组和修复的不一致性。虽然该假说没有完全解释病毒的起源，但解释了复制系统的多样性，以及某些类型的 DNA 病毒和 RNA 病毒具有非常原始的结构相似性。

制。再如，Amer Qureshi 设计了一个锤头状核酶(图 11-9)，它以 HIV 感染宿主细胞必需的辅助受体，即趋化因子受体——CXCR4 的 mRNA，作为目标 RNA，结果很好地抑制了 HIV 对宿主细胞的感染。

为提高核酶的稳定性和切割效率，可对设计合成好的核酶在体外进行化学修饰，修饰位点通常在 2′-OH。实验证明，经过修饰的核酶在体外的活性可提高 6 倍，稳定性提高 2 倍。此外，现在已出现了可对靶 RNA 分子上多个位点进行切割的第二代核酶，它们可以大大提高核酶的剪切效率。

科学故事——核酶的发现

核酶的发现主要归功于美国科学家 Thomas Cech。1978 年，Cech 在美国科罗拉多大学建立了自己的实验室以后，决定研究某一个基因的结构和功能。他选择了四膜虫的 rRNA 基因即 rDNA 作为研究对象。像四膜虫一样的纤毛类原生动物，转录发生在大核(macronucleus)中。据估计，大核含有 1 万个拷贝的 rRNA 的基因。

吸引 Cech 注意的倒不是在大核中含量丰富的 rRNA，而是核中现成的各种转录必需的蛋白质因子。实验的起始步骤并不复杂，只是利用四膜虫的大核制备无细胞转录系统，并看看能否利用制备好的系统进行体外转录。为此他在系统中，加入了 RNA 聚合酶、rDNA、4 种 NTPs(有 1 种被放射性同位素标记)、标准的转录缓冲液以及用来抑制 RNA 聚合酶Ⅱ和Ⅲ活性的 α - 鹅膏蕈碱(目的是只允许 rDNA 基因的转录)(参看第三十六章“DNA 转录”)。

Cech 得到的实验结果一方面正如预测的一样，另一方面却出乎意料：他观测到了一种 26S RNA 的合成，这与预期的 rRNA 前体的大小(26S)是一致的；然而，他也观测到一种较小的 9S RNA 的合成，并且这种小 RNA 的量随着反应时间的推移而不断积累。起初，他认为 9S RNA 一定是 rRNA 前体后加工反应的副产物，比如是 5′ - 外部转录间隔序列(external transcribed spacer，ETS)或对应于 17S 和 5.8S rRNA 之间的间隔序列。

为了弄清楚 9S RNA，他安排他的学生 Art Zaug 去鉴定它。令他们吃惊的是，9S RNA 居然是当时已经知道的位于 26S 基因内部的内含子。这的确是一个激动人心的发现！这意味着 Cech 的体外转录系统不但能转录出含有内含子的基因，而且转录产物在体外还能剪接出内含子。于是，他们以此建立了一个简单易行的研究剪接机制的系统。根据每一个细胞约含有 1 万个拷贝的 rRNA 基因的事实，Cech 推想，如果每一个基因都先转录再剪接，并以每秒钟 1 拷贝 rRNA 的速率，他研究的大核里就应该含有丰富的剪接酶。

为了确定反应系统的性质，他们首先得确定 rRNA 前体能合成但剪接不能发生的条件，因为这样才可以分离到 rRNA 前体作为底物，以便在试管里研究剪接反应。他们设置的反应很简单：在一个试管里，将 rRNA 前体与核抽取物一起保温，使用的缓冲溶液与体外转录反应相同，以观察剪接反应；在另一个进行对照实验的试管中，省掉了核抽取物。

然而结果又是一个惊喜：两个试管里都发生了剪接反应！面对意外的结果，Cech 当时对 Art 说：“如果你在制备对照样品的时候没有犯什么错误的话，这结果看起来很是鼓舞人心。”面对意外的结果，做导师的对自己的学生会说过多少次这样的话！而学生出错的机会又有多大！但这一次 Art Zaug 是无辜的。

在确认结果无误以后，他们下一步就是测定剪接出来的内含子序列，以确定内含子是不是被正确地剪切掉了。当测序结果出来以后，他们发现序列与已知的内含子序列完全匹配，但在内含子的 5′端多出了 1 个鸟苷酸。考虑到以前测定过的序列可能有错，他们给最初测定内含子序列的 Joe Gall 实验室打了电话，询问是不是遗漏了 5′端 1 个鸟苷酸。但 Joe Gall 实验室否认有任何错误，且 5′剪接点也不可能有鸟苷酸。

与此同时，Cech 在努力细究体外剪接反应所必需的成分。他发现，去除 4 种 NTPs 中的 3 种 NTPs (ATP、UTP 和 CTP)对反应无任何影响，但是 GTP 却是绝对必需的成分。于是，Cech 疑问：5′端多出了鸟

苷酸与体外剪接反应必需鸟苷酸难道只是一种巧合吗？

为此，Cech 提出假设，认为剪接反应中需要的 GTP 被添加到内含子的 5′端。验证这个假设的实验并不困难，只需将[^{32}P]- 标记的 GTP 与非标记的 rRNA 前体在一起保温，观察内含子的标记与它的切除是不是同时发生。然而，实验的结果却让 Cech 认为这是他曾做过的最奇怪的实验：一方面，实验的成功基于对使用的实验系统有一定的认识而做出的简单预测；另一方面，他难以相信，将一个核苷酸与经过酚抽取和蛋白酶处理过的 RNA 简单地混在一起，就可能导致一个共价键的形成。按照 Cech 自己的说法，当时他不想在自己的研究生和同事面前因为可能的失败而出丑，因此他所做的一切都是悄悄地进行。

结果可想而知，实验是成功的。在经过进一步确认和鉴定以后，他们相信，内含子的剪接不需要任何蛋白质，只需要内含子本身和鸟苷酸或鸟苷。

后来，Cech 首次使用了 ribozyme 来描述他发现的具有催化活性的 RNA 分子，但从严格的意义来说，他发现的自我剪接的内含子并不具有真正的催化活性，这是因为它催化的是自身而不是其他分子的切除。不久以后，美国的另外一位科学家 Sidney Altman 发现，大肠杆菌的核糖核酸酶 P 才是一种真正的核酶。Cech 和 Altman 因发现 RNA 自身具有酶的催化功能而获得 1989 年诺贝尔化学奖。

e11-4 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Ribozyme（维基百科有关核酶的内容）
2. http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/R/Ribozymes.html（一些天然核酶的介绍）
3. http://www.panspermia.org/rnaworld.htm（内有“RNA 世界”假说的详细描述和相关证据）
4. http://www.proteopedia.org/wiki/index.php/Ribozyme（维基百科提供的有关各种核酶的资料）

参考文献：

1. Wochner A, et al. Ribozyme-catalyzed transcription of an active ribozyme. *Science*, 332:209–212, 2011.
2. Turk R M, Chumachenko N V, Yarus M. Multiple translational products from a five-nucleotide ribozyme. *Proc. Natl Acad Sci USA*, Published online February 22, 2010. DOI:10.1073/pnas.0912895107
3. McManus S A, Li Y F. The structural diversity of deoxyribozymes. *Molecules*, 15:6 269–6 284, 2010.
4. Breaker R R, et al. Control of gene expression by a natural metabolite-responsive ribozyme. *Nature*, 428: 281–286, 2004.
5. Nakano S, et al. General acid-base catalysis in the mechanism of a hepatitis delta virus ribozyme. *Science*, 287:1 493–1 497, 2000.
6. Guerrier T C, et al. The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. *Cell*, 35:849–857, 1983.
7. Cech T R, et al. In vitro splicing of the ribosomal RNA precursor of Tetrahymena: involvement of a guanosine nucleotide in the excision of the intervening sequence. *Cell*, 27:487–496, 1981.

第十二章 酶活性的调节

具有高效催化性能的酶就像高速行驶的赛车一样，速率越快，危险就越大，因此需要有专门的变速装置——“油门”和“刹车”，以便能及时对速率进行调节。生物体内的酶所使用的“油门”通常是激活剂，而“刹车”是抑制剂，除此以外，还有其他的变速装置。有了这些装置，才可以保证一种酶在正确的时间、正确的地点具有正确的活性。如果调节机制失灵，轻者会导致细胞功能的紊乱，重者就会诱发疾病的发生，甚至导致细胞或机体的死亡。

酶活性的调节主要有两种手段，一种是通过改变酶浓度，即以“量变”的方式进行；另一种通过改变已有的酶的活性，即以“质变”的方式进行。

本章主要介绍酶活性的“质变”调节，特别是其中的别构调节和共价修饰调节。

第一节 酶的“量变”

一、酶的“量变”和“质变”的主要差别

酶的“量变”和“质变”这两种调节手段在调节速率、能耗、酶活性的限速因素和作用持续时间等方面都有显著的差别（表 12–1）。有的酶会同时使用这两种手段，如参与胆固醇合成的限速酶——HMG–CoA 还原酶（参看第三十章“胆固醇代谢”），但一般只以其中的一种方式为主。

表 12–1 酶的“量变”和“质变”的比较

	量变	质变
调节速率	慢，几小时至几天	快速，几秒钟至几分钟
能耗	高（通常涉及酶基因的表达，因此需要消耗大量的 ATP）	低（除非使用抑制蛋白，因为在解除抑制的时候，通常需要将抑制蛋白水解）
决定酶最高活性的主要因素	酶合成与水解的相对速率	已有的酶浓度
活性变化的持续时间	长	短

改变酶量的方式有两种，一种是通过同工酶（isozymes），另外一种是通过控制酶基因的表达和酶分子的降解。

二、同工酶

COO^- | HO—C—H | CH_3 + NAD^+

L-乳酸

心脏(H_4) ⇅ 骨骼肌(M_4)

COO^- | C=O | CH_3 + NADH + H^+

丙酮酸

图 12–1 乳酸脱氢酶催化的反应

同工酶是指催化相同的反应但性质不同（V_{max} 和 / 或 K_m 不同）的酶。它们可能以不同的量出现在同一种动物不同的组织或器官，也可能出现在同一个细胞但位于不同的细胞器。例如，高等动物的乳酸脱氢酶（lactate dehydrogenase，LDH）有 M_4、M_3H、M_2H_2、MH_3 和 H_4 五种形式（图 12–1）。M_4 由四个 M 亚基组成，主要存在于骨骼肌。H_4 由四个 H 亚基组成，主要存在于心脏。又如，在真核细胞的细胞质和线粒体基质，各有一种不同的苹果酸脱氢酶（malate dehydrogenase）。

同工酶的差异可能是在一级结构上，由不同的基因编码或者由同一个基因编码但经过不同的转录后加工造成的，也可能是在四级结构上，比如构成 LDH 的两种亚基的不同组合，还可能是在翻译的后加工上。

同工酶之间在不同组织或不同亚细胞空间内有相对丰度的差别，或者一种同工酶只能在某种特定的细胞里表达，可允许细胞根据胞内特定的生理状况而对酶活性进行调节。例如，骨骼肌细胞中的

M4-LDH 对丙酮酸的 K_m 比对乳酸的低，由于骨骼肌容易缺氧，因此 LDH 在肌肉细胞内的主要功能是促进乳酸的形成，以便糖酵解能够正常地进行；而心肌细胞是不允许缺氧的，其内的 H4-LDH 对乳酸的 K_m 更低，因此心脏的 LDH 的主要功能是促进乳酸在心肌细胞内的分解。

此外，同工酶之间还可以在质变的调控方式上存在差异，这就进一步增加了细胞对酶活性调控的灵活性。

三、酶的合成和降解

控制酶浓度的手段主要有调节酶基因的表达（参看第四十二章“原核生物的基因表达调控”和第四十三章“真核生物的基因表达调控”），此外还有受控的蛋白质的酶促降解，特别是泛素介导的依赖于蛋白酶体的蛋白质水解（参看第三十九章“蛋白质的生物合成及其在细胞内的降解”）。

第二节　酶的“质变”

酶“质变”的方式包括别构调节、共价修饰、水解激活、调节蛋白的结合和解离以及单体的聚合和解离。

一、别构调节

别构调节也称为变构调节，这是所有酶活性调节方式中最快的一种。其原理在于一些酶除了活性中心以外，还含有别构中心。别构中心能够结合一些特殊的配体（有时为底物）。当别构中心结合配体以后，酶构象发生改变，从而影响到活性中心与底物的亲和力或催化能力，并最终导致酶活性发生变化。能够进行别构调节的酶称为别构酶，与别构中心结合并调节酶活性的配体称为别构效应物，其中起抑制作用的别构效应物称为别构抑制剂或负别构效应物，起激活作用的别构效应物称为别构激活剂或正别构效应物。由底物作为别构效应物产生的别构效应称为同促效应，反之称为异促效应。许多别构酶具有多个别构中心，能够与不同的别构效应物结合，例如大肠杆菌的谷氨酰胺合成酶的别构效应物至少有 8 种。

关于别构酶的主要性质可以参看第九章第五节“别构酶的动力学”。别构调节最多出现在代谢途径中的反馈抑制（feedback inhibition），它是指一条代谢途径（通常是合成代谢途径）的终产物作为别构抑制剂，抑制位于其上游的限速酶的活性，从而关闭自身合成的一种调节方式（图 12-2），因此也被称为终产物抑制（end-product inhibition）。反馈抑制通常的情形是一条代谢途径的最后一步反应的产物，去抑制第一步反应的酶［图 12-2(1) 中的 D 抑制 E_1］，但如果代谢途径途中出现分支［图 12-2(2) 中 C 开始］，则每一分支途径的终产物（F 或 Y）一方面反馈抑制分支点的酶活性［图 12-2(2) 中的 F 抑制 E_3 和 Y 抑制 E_5］，另外一方面还各自部分抑制第一步反应的酶（E_1）。只有当两个终产物（F 和 Y）同时与 E_1 结合的时候，E_1 才会被完全抑制，这样可以保证在 F 过量、Y 不足的时候，只会关闭 F 的合成而不会关闭 Y 的合成，或者在 Y 过量、F 不足的时候，只会关闭 Y 的合成而不会关闭 F 的合成。

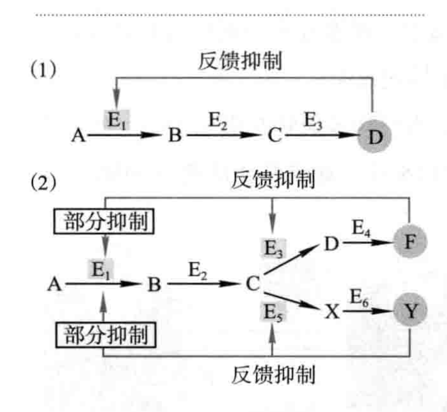


图 12-2　酶活性的反馈抑制

反馈抑制使得细胞能够对胞内一些重要的代谢物浓度的变化迅速做出反应，这对于维持细胞内不同代谢物浓度的平衡至关重要。当一条代谢途径的终产物的量积累到一定程度，已能满足机体需要的时候，过量的终产物就反馈抑制第一个限速酶的活性而关闭自身的合成。这一方面可以阻止终产物的进一步堆积而可能产生的毒性，另一方面还可以将能量用于其他代谢途径。一旦终产物因消

Quiz1 生物体一般使用负反馈机制，以维持内环境的稳定，但生物也有使用正反馈机制的时候。请给出使用正反馈的一个例子。

耗而降低到一定水平时，即与限速酶解离，于是别构抑制解除，合成它们的代谢途径得以重新开放。

除了反馈抑制以外，机体有时候还会使用前馈激活（feed-forward activation）和底物激活（substrate activation）这两种方式进行别构调节。前馈激活是指一条代谢途径位于上游的代谢物作为别构激活剂，激活下游的限速酶，而底物激活属于同促调节（homotropic regulation），它具有两个方面的含义，其一是指底物与酶活性中心结合产生的正协同效应，其二是指底物与酶的别构中心结合，激活酶的活性。前馈激活的一个例子出现在糖酵解，糖酵解第三步反应产生的1,6-二磷酸果糖（fructose-1,6-bisphosphate）作为别构激活剂，激活催化最后一步反应的丙酮酸激酶（pyruvate kinase）的活性。这种调节的好处是让已经进入糖酵解途径的底物尽快被糖酵解氧化分解（详见第二十二章“糖酵解”）。与正协同效应相关的底物激活很常见，绝大多数别构酶都具有这个性质。但使用第二种形式底物激活较为罕见，它可能出现在机体对一些外来毒物的解毒反应中。这样可以及时快速地解除毒物的毒性，尽可能降低它们对机体造成的毒害。

有两种模型可以解释别构酶的别构效应和与底物结合的协同效应，一种是齐变模型，另一种是序变模型。

（一）齐变模型

齐变模型也称为对称模型（the symmetry model），由Monod、Wyman和Changeux于1965年提出，因此也被简称为MWC模型。它基于别构酶由多个亚基组成的事实，认为构成别构酶的亚基能够以两种不同的构象形式存在，一种构象为松弛态或R态，另外一种构象为紧张态或T态。在一个特定的酶分子内部，构成亚基之间的相互作用致使每一个酶分子的每一个亚基在某一个时候采取同一种构象，即要么都是R态，要么都是T态，没有R态亚基和T态亚基的杂合体。

在溶液中，两种构象可以相互转变，并处于动态的平衡中，但转变的方式为齐变，即构成它们的亚基要么一齐从R态变成T态，要么一齐从T态变成R态，如左上图所示。

如果溶液中无任何配体（底物或别构效应物），平衡主要偏向右边，即几乎所有的酶都处于T态。虽然R态和T态上的活性中心都能结合底物，但是R态对底物有更高的亲和性。

如果将少量的底物加入溶液，R态酶由于对底物有更高的亲和力，就更容易与底物结合。当一个底物分子与R态酶结合以后，就形成左下图所示平衡。

于是，底物与R态酶的结合等于是从溶液中移走了一些游离的R态酶，根据化学平衡理论，平衡会被拉向左边，结果导致溶液中有更多的R态酶，而T态酶相应减少。既然R态酶对底物有更高的亲和性，也就提高了酶与底物总的亲和性，由此产生了底物结合的正协同效应。

使用齐变模型也容易解释别构效应物的作用原理（图12-3）：当别构效应物在别构中心与酶结合以后，诱导酶的构象发生变化，从而打破了R态酶和T态酶之间的平衡。如果是激活剂，就更容易与R态酶结合，从而像底物一样，将平衡拉向右边，使更多的酶转变为R态，最终产生激活的效果；如果是抑制剂，则更容易与T态酶结合，以致将平衡拉向左边，致使更多的酶变成T态酶，最终产生抑制。

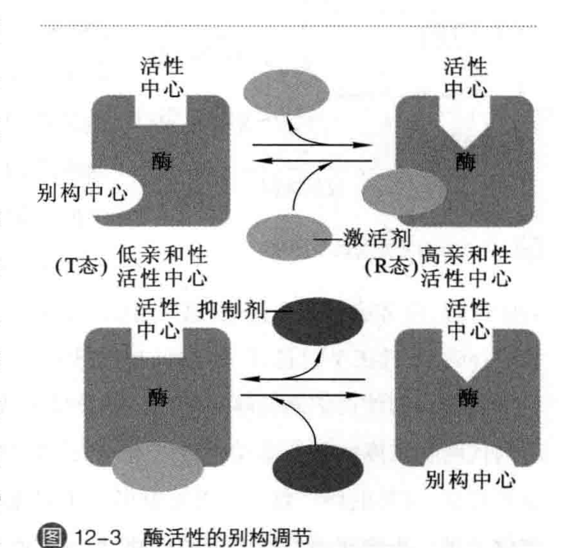


图 12-3 酶活性的别构调节

总之，激活物的存在意味着溶液中有更多的酶处于R态，这就省掉了一些底物去诱发T态向R态转变，因此底物结合的正协同性在表面上是下降了。如果存在大量的激活剂，溶液内几乎所有的酶就都处于R态，这就意味着底物结合不再产生协同性，这时动力学曲线从S形也就变成了双曲线。而抑制剂的作用则正好相反，它会导致溶液中更多的酶处于T态，这意味着需要更高浓度的底物才能实现同样的



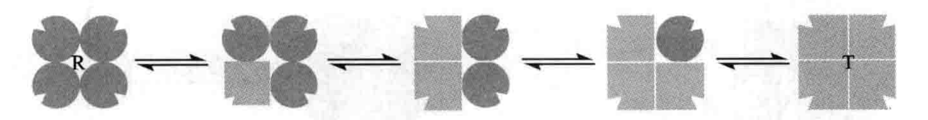
正协同效应，S 型曲线就变得更明显了。

虽然齐变模型能很好地解释别构酶的一些性质，但对于某些别构酶来说，可能过于简单了，并且它也不能解释某些别构酶的负协同性。

（二）序变模型

该模型由 Koshland、Nemethyl 和 Filmer 于 1966 年提出，因此也被简称为 KNF 模型。与齐变模型最大的不同在于序变模型接受了混合酶存在的可能性，即同一个酶分子既有 R 亚基，又有 T 亚基，也就是溶液中的 R 态酶（R4）和 T 态酶（T4）之间存在多种混合体（R3T1、R2T2、R1T3），各种状态的酶处于动态平衡之中：

Quiz2 为什么齐变模型不能解释别构酶对底物的负协同性？



此外，序变模型还肯定了底物对酶构象有更直接的影响。在没有底物时，酶差不多都以 T 态存在，这时活性中心的构象不是与酶结合的最佳构象。一旦底物进入活性中心，活性中心构象发生变化，致使底物与酶结合得更加帖服，这就是“诱导契合”。“诱导契合”导致与底物结合的亚基从 T 态转变成 R 态。

序变模型还认为相邻亚基之间存在相互作用，并且这种相互作用可以影响到其他亚基的构象状态。仍然以具有底物正协同性的别构酶为例加以说明，当 T 态酶上的一个亚基因为底物的结合从 T 态变为 R 态以后，该亚基会促进其他的亚基向 R 态转变，致使其他亚基能够以更高的亲和性与底物结合。

使用序变模型也很容易解释别构效应物的作用原理：激活剂仅仅在别构中心与酶结合，通过与底物一样的方式促进 T 亚基变为 R 亚基，而抑制剂与酶的结合使酶的构象变得更为僵硬，很难通过诱导契合从 T 态变为 R 态。

使用序变模型还有一个好处，就是很容易解释某些酶具有的底物结合的负协同性这一现象，但用齐变模型难以解释，这是因为齐变模型完全依赖于化学平衡理论将 R/T 平衡拉向高亲和性的 R 态一边，而在反应系统之中加入底物，底物总是优先与高亲和力的 R 态结合，将平衡从 T 态拉向 R 态，而不是将 R 态拉向 T 态。然而利用序变模型就很容易解释了，只要假定一个底物与一个亚基结合以后，通过亚基之间的相互作用，导致其他的亚基更难从 T 态变成 R 态。

至于两种模型孰优孰劣，单从负协同性来看，序变模型似乎是正确的。但从另外一个方面来看，在序变模型中，R 亚基和 T 亚基以混合的方式存在会导致更为复杂的平衡。而且，使用快速反应技术（fast reaction techniques）对某些别构酶的研究表明，这些酶具有非常简单的平衡样式，此时用齐变模型解释似乎更为合理。

（三）别构酶实例——氨甲酰转移酶

天冬氨酸转氨甲酰酶（aspartate transcarbamylase，ATCase）或氨甲酰转移酶（carbamoyltransferase）是大肠杆菌嘧啶核苷酸从头合成途径中的限速酶（详见第三十二章“核苷酸代谢”），它催化氨甲酰磷酸和 Asp 形成 *N*- 氨甲酰天冬氨酸和无机磷酸，其活性受到严格的调控。

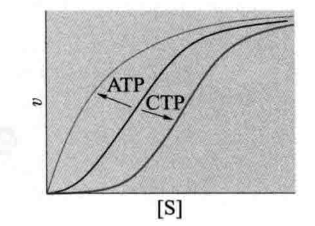


图 12-4　ATP 和 CTP 对氨甲酰转移酶活性的调节

对大肠杆菌 ATCase 的动力学研究表明，其动力学曲线为 S 形，活性受到嘧啶合成的终产物 CTP 的反馈抑制，但受到嘌呤核苷酸 ATP 的激活（图 12-4）。S 形曲线表明该酶与底物结合具有正协同性。

为了搞清楚大肠杆菌 ATCase 的正协同性以及 ATP 和 CTP 对酶活性的调节机制，人们对它的结构作了大量的研究，结果表明，大肠杆菌的 ATCase 全酶由 12 个亚基组成，包括 6 个大的催化亚基（C 亚基）和 6 个小的调节亚基（R 亚基）。其中每 3 个 C 亚基构成 1 个催化三聚体，每 2 个 R 亚基构成 1 个调节二聚体，这样 1 个全酶分子就是由 2 个催化三聚体和 3 个调节二聚体组成。它们按照一定的方式排列在一起，其中两个催化三聚体各构成酶分子的两个主要的面，而调节二聚体构成酶的几个角呈三角形（图 12-5）。

每一个催化三聚体单位含有 3 个活性中心。对在酶竞争性抑制剂——*N*- 膦乙酰基 -L- 天冬氨酸（*N*-phosphonacetyl-L-aspartate，PALA）（图 12-6）存在下得到的 ATCase 晶体进行 X 射线衍射分析，

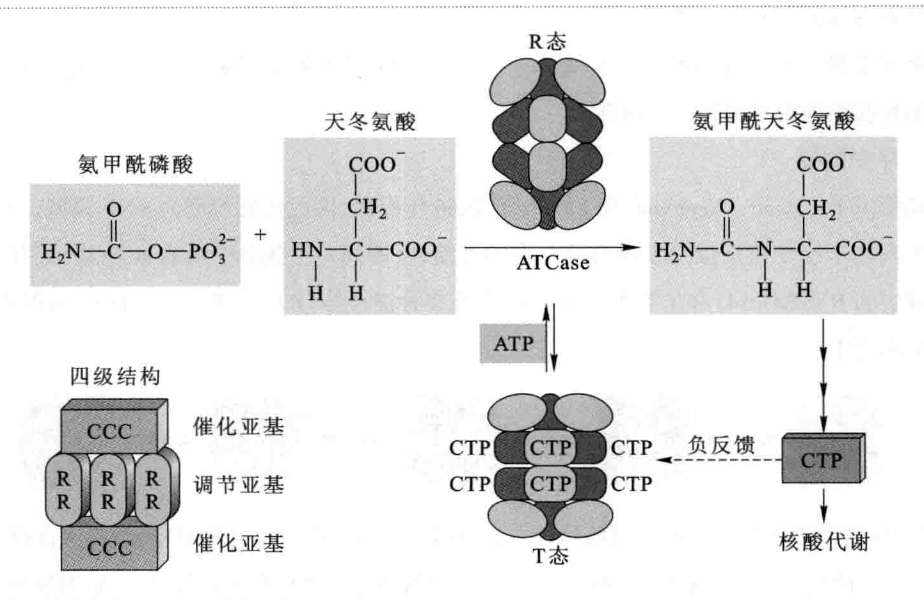


图 12-5　氨甲酰转移酶的结构及其催化的反应

结果表明组成活性中心的大多数残基属于同一个亚基，但少数重要的残基属于相邻的亚基，PALA 的结合位点就位于一对 C 亚基的界面。

PALA

图 12-6　*N*-膦乙酰基-L-天冬氨酸的化学结构

对在 CTP 存在下得到的 ATCase 晶体的研究表明，CTP 与调节亚基结合，它离最近的活性中心有 5 nm。由于催化三聚体和调节二聚体是通过与 Cys 残基结合的 Zn^{2+} 而聚合在一起的，使用巯基试剂——对羟基汞苯甲酸（p-Hydroxymercuribenzoate）与 ATCase 上的 Cys 巯基反应，可以人为地将它们分开。

单独的催化三聚体保留催化活性，但无别构效应，也无底物结合的正协同性；单独的调节二聚体无催化活性，但仍然能够与酶的别构抑制剂（CTP）和激活剂（ATP）结合。分离开的催化三聚体和调节二聚体在体外能够重组成天然的酶，并恢复原来的别构性质和与底物结合的正协同性。

Quiz3　若是将一个催化三聚体的三个亚基分开，你认为单个亚基还具有催化活性吗？为什么？

比较有无 PALA 结合时的 ATCase 晶体结构发现，一旦 PALA 结合，ATCase 的构象就发生很大的变化：首先 2 个催化三聚体分开 1.2 nm 的距离，并旋转了 10°；其次，为配合催化三聚体的运动，调节二聚体也旋转了 15°（图 12-7）。

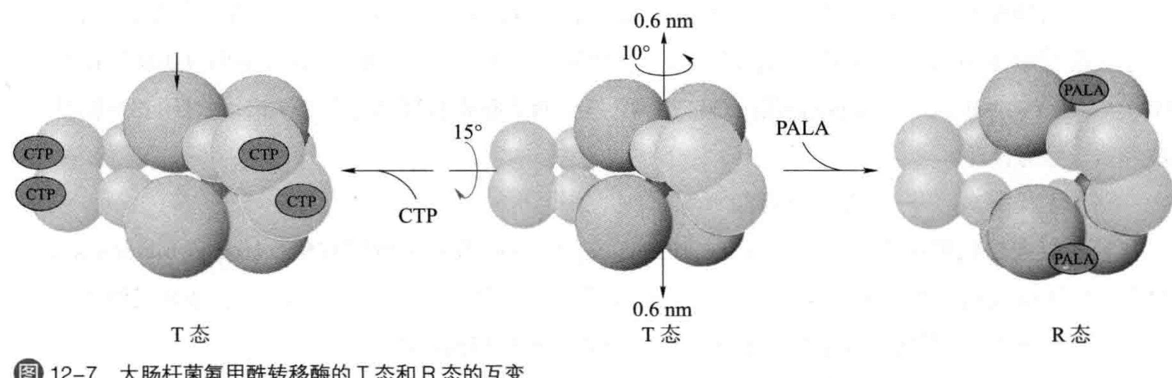


图 12-7　大肠杆菌氨甲酰转移酶的 T 态和 R 态的互变

综合上面的研究数据，并结合序变模型或齐变模型，就能很容易解释 ATCase 对底物的正协同性、CTP 对酶活性的抑制以及 ATP 对酶活性的激活。

使用齐变模型对 ATCase 的底物正协同性的解释：处于 R 态酶上的一个活性中心与底物结合以后，溶液中有更多的酶从 T 态转变成 R 态，使得后面的底物更容易与活性中心结合；使用序变模型对 ATCase 的底物正协同性的解释是：处于 T 态构象的一个活性中心与底物结合以后变成了 R 态，同时相邻的活性中心受到该亚基的影响也变成 R 态，或者更容易变成 R 态，以致后面的底物与活性中心的结合更为容易。

对于 CTP 或 ATP 对 ATCase 活性调节的解释是：CTP 和 ATP 都是酶的别构效应物，前者与 T 态

的酶(调节亚基)结合以后,稳定 T 态,致使更多的酶从 R 态转变为 T 态,从而充当别构抑制剂,而 ATP 与酶结合以后,所引起的结果与 CTP 正好相反,使溶液中更多的酶从 T 态转变成 R 态,因而是一种别构激活剂。

ATP 和 CTP 作为别构效应物对 ATCase 活性调节的生理意义在于:高浓度 ATP 的存在是细胞内高浓度嘌呤核苷酸的信号,它作为别构激活剂激活 ATCase 活性,有利于平衡胞内的嘌呤核苷酸库和嘧啶核苷酸库;CTP 作为嘧啶核苷酸合成的终产物之一,反馈抑制 ATCase 活性也有利于维持两类核苷酸库的平衡。

二、共价修饰调节

酶的共价修饰调节是指酶活性因其分子内的某些氨基酸残基发生共价修饰而发生变化的过程。这是由修饰酶(modifying enzymes)和去修饰酶(de-modifying enzymes)共同构成的一种可逆的环式调节系统。由修饰酶催化的共价修饰的方式有(表 12-2):磷酸化、腺苷酸化、尿苷酸化、ADP- 核糖基化、甲基化和形成二硫键。其中磷酸化是最为常见的形式,主要发生在真核细胞。腺苷酸化和尿苷酸化很少见,仅发现在某些原核细胞。形成二硫键见于植物(参看第二十六章“光合作用”第六节“卡尔文循环的调节”)。如果是磷酸化,则修饰酶是蛋白质激酶,去修饰酶是磷蛋白磷酸酶。

表 12-2 酶共价修饰的几种形式

修饰方式	基团供体	基团受体	分布
磷酸化	ATP	真核通常为 S、T 和 Y;原核通常为 H 和 D	主要是真核细胞,少数为原核细胞
腺苷酸化	ATP	Y	原核细胞
尿苷酸化	UTP	Y	原核细胞
ADP- 核糖基化	NAD^+	R、E、K、C 和 H	原核细胞和真核细胞
甲基化	*S*- 腺苷甲硫氨酸	D、E、K、H 和 Q	原核细胞和真核细胞
形成二硫键	硫氧还蛋白	C	植物

这种调节方式比别构调节要慢,但与其他调控机制相比,它能够对更多的效应物刺激产生反应,在调控的样式上具有更大的灵活性,并且能够对效应物浓度的变化做出更强烈的反应。这是因为参与共价修饰与去共价修饰的酶本身又受到别构调控,一种修饰酶的别构效应物浓度发生的细微变化可导致被修饰的靶酶在活性上发生较大的变化。

以蛋白质的磷酸化 / 去磷酸化为例(图 12-8),在蛋白质激酶的催化下,许多酶上的羟基氨基酸,例如 Ser、Thr 或 Tyr 上的羟基接受 ATP 分子上的 γ- 磷酸根而被磷酸化修饰,被修饰的酶活性因构象变化发生改变。这种改变是可逆的,在磷蛋白磷酸酶催化下,磷酸根被水解,酶又恢复到原来的状态。

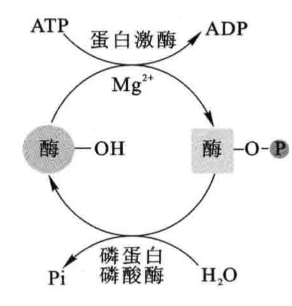


图 12-8 蛋白质的可逆磷酸化

磷酸化对蛋白质的影响包括:①增加了 2 个负电荷,影响其静电作用;②磷酸化基团可形成 3 个氢键;③较大的自由能变化,贮存在磷蛋白上;④磷酸化和去磷酸化可在很长的时段内发生,间隔的时间是可以调节的;⑤产生级联放大。一种激酶的底物可能是另外一种激酶,如此作用具有放大效应。

必须注意的是,某些酶因磷酸化修饰从无活性变为有活性,有些酶正好相反。磷酸化形式才有活性的酶有:糖原磷酸化酶和对激素敏感的脂肪酶;去磷酸化形式有活性的酶有:糖原合酶、丙酮酸激酶、丙酮酸脱氢酶、乙酰 -CoA 羧化酶、HMGCoA 还原酶。

真核细胞内含有许多蛋白质激酶的现象说明了蛋白质磷酸化的重要性。据估计,人类基因组编码了 1 000 种以上不同的蛋白质激酶。各种激酶的特异性彼此有别,按照接受磷酸根的氨基酸残基的性质,蛋白质激酶可分为丝氨酸 / 苏氨酸蛋白质激酶、酪氨酸蛋白质激酶和双功能激酶(既可修饰 Ser/Thr,还可以修饰 Tyr)。微生物体内还存在一种组氨酸蛋白质激酶。与蛋白质激酶相比,催化磷蛋白去磷酸化的磷酸酶的种类只有几百种,且作用的特异性要广,进化的保守性也比较强。

Quiz4 还有哪些氨基酸的侧链可以被磷酸化修饰?

三、水解激活

一些酶(主要是蛋白酶)在体内以无活性的酶原形式被合成,需要通过水解(由其他蛋白酶催化或自我催化)去除一些氨基酸序列以后才会具有活性,这种调节酶活性的方式称为水解激活。

与共价修饰一样,水解激活也是一种全或无的调节方式,酶原状态没有活性,但与共价修饰不同的是,它是不可逆的,即一旦被激活就不可能再回到原来的非活性的酶原状态。

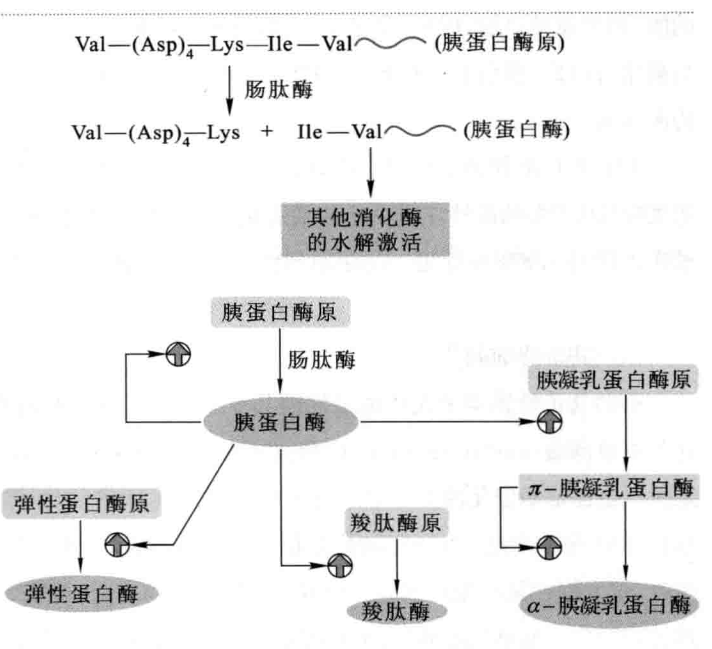


图 12-9 胰蛋白酶等几种蛋白酶的水解激活

通过这种机制调节酶活性的主要是一些消化酶,如胃蛋白酶、胰蛋白酶、胰凝乳蛋白酶、羧肽酶和弹性蛋白酶。这些消化酶以酶原的形式被分泌到消化道,然后再被水解激活。如果它们在细胞内提前激活,就会导致细胞自溶。医学研究表明,消化酶原在细胞内的提前活化会导致急性胰腺炎等疾病的发生。此外,几种与凝血有关的凝血因子、参与细胞凋亡的胱天蛋白酶以及补体激活途径中的某些成分也需要水解才能激活。

以胃蛋白酶为例,其酶原多出 44 个氨基酸残基。在由胃主细胞分泌到胃腔以后,胃蛋白酶原受到胃酸的作用,构象发生变化而发生自切割,在丢掉 N 端 44 个残基以后被激活。先行激活的胃蛋白酶再作用其他还没有激活的酶原,产生更多的有活性的胃蛋白酶。

Quiz5 图 12-9 中的哪一种突变蛋白酶突变造成的危害更大?为什么?

如果是胰蛋白酶原,它在从胰腺分泌到肠腔后,就需要肠肽酶(enteropeptidase)或已激活的胰蛋白酶在酶原内部的 Lys—Ile 处将其切开,去除 N 端一段六肽序列以后才能有活性(图 12-9)。

胰凝乳蛋白酶的激活略微复杂(图 12-10):其前体只有一条肽链组成,共有 245 个氨基酸残基,内有 5 个链内二硫键。首先它在胰蛋白酶作用下,位于 Arg15—Ile16 之间的肽键被水解,转变为 π-胰凝乳蛋白酶,但水解产生的 N 端十五肽和 C 端多肽仍然通过 Cys1 和 Cys122 之间的二硫键连在一起。游离出来的 Ile16 的 α-氨基因质子化成为带正电荷的基团,与紧靠活性中心的 Asp195 架起盐桥,从而导致活性中心的移动,使 Gly193 和 Ser195 处于合适的位置,能够与四面体过渡态上的氧阴离子形

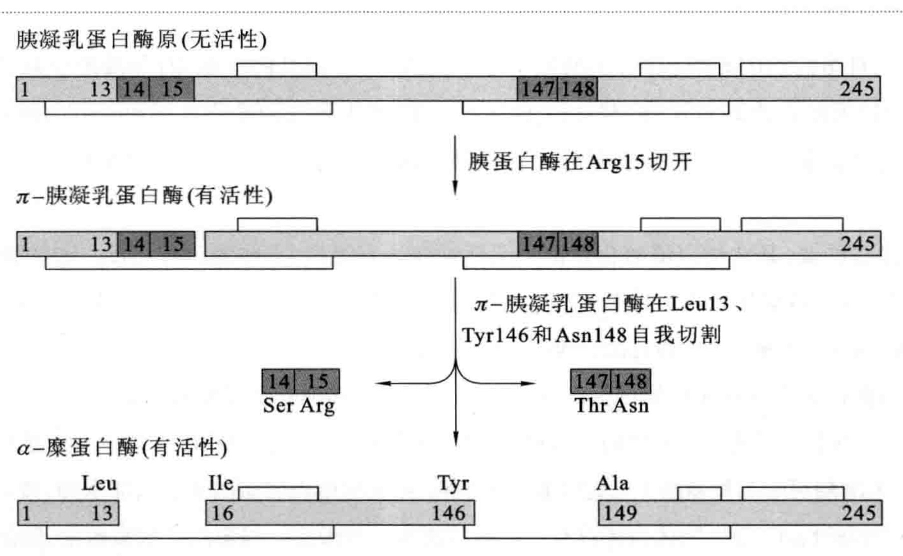


图 12-10 胰凝乳蛋白酶的水解激活

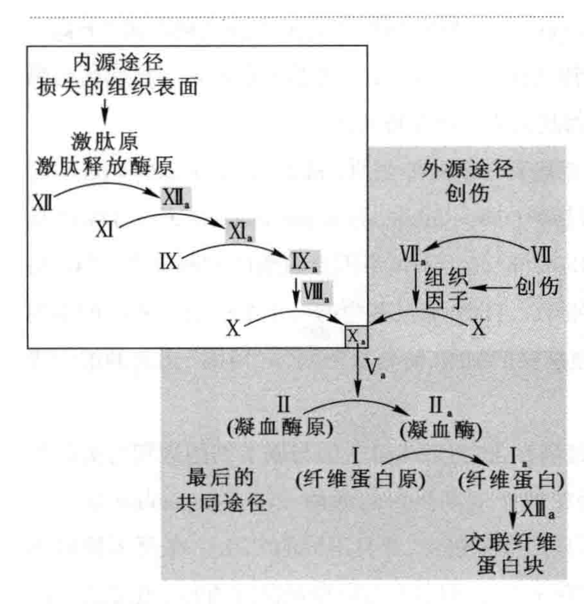


图 12-11　某些凝血因子的水解激活

成氢键(参看上一章有个胰凝乳蛋白酶的作用机理)。至此,π-胰凝乳蛋白酶已有活性,但稳定性不高。随后,π-胰凝乳蛋白酶通过相互催化,将 Leu13—Ser14、Tyr146—Thr147 和 Asn148—Ala149 之间的三个肽键水解,释放出两个二肽(Ser14—Arg15 和 Thr147—Asn148),最终产生稳定的有活性的α-胰凝乳蛋白酶。

除了消化酶以外,绝大多数参与血液凝固的凝血因子也是被水解激活的(图 12-11),这些凝血因子包括Ⅻ、Ⅺ、Ⅸ、Ⅹ、Ⅶ、Ⅱ、V 和Ⅷ,其中前六种属于丝氨酸蛋白酶,而凝血因子Ⅱ就是凝血酶(thrombin)。这些凝血因子在特定的条件下(内源途经或外源途经)依次被水解激活,构成凝血的级联反应,最后纤维蛋白原(fibrinogen)被凝血酶水解激活,丢掉纤维蛋白肽(fibrinopeptide)A 和 B 后变成纤维蛋白(fibrin)。纤维蛋白在释放出这两种小肽以后,暴露出相互结合的位点,进而彼此聚合成纤维状的血凝块。随后纤维蛋白之间在 Glu 和 Lys 残基处形成共价交联,在凝血因子ⅩⅢ的作用下形成更稳定的聚合体。然而,随着伤口的愈合,血凝块必须被清除,否则会诱发中风或心脏病。此过程主要由纤溶酶(plasmin)催化,该酶也是以酶原的形式被合成,受组织纤溶酶激活物(tissue plasminogen activator,tPA)的水解激活。

Quiz6 胰凝乳蛋白酶的活性受胰蛋白酶的激活,而不能反过来,即让胰蛋白酶活性受胰凝乳蛋白酶的激活。这是为什么?

血友病(hemophilia)与凝血因子Ⅷ或Ⅸ的基因缺陷有关。患者一旦出现了伤口,便会血流不止。唯一有效的治疗方法是输入外源的凝血因子Ⅷ或Ⅸ。此外,维生素 K 的缺乏也会影响到凝血过程,这是因为许多凝血因子在翻译后需要对 Glu 残基进行特殊的γ-羧基化修饰,而这种修饰反应需要维生素 K 作为辅酶(参看第十四章“维生素与辅酶”)。

四、调节蛋白的激活或抑制

某些蛋白质也能够作为配体,与特定的酶结合从而调节被结合的酶的活性,这些调节酶活性的蛋白质称为调节蛋白。其中,激活酶活性的调节蛋白称为激活蛋白,抑制酶活性的调节蛋白称为抑制蛋白。抑制蛋白通常结合在酶的活性中心,通过阻止底物与活性中心的结合来达到抑制的效果。

抑制蛋白中最常见的一类是丝氨酸蛋白酶抑制剂(serine protease inhibitor,Serpin),它们专门与丝氨酸蛋白酶结合,并抑制丝氨酸蛋白酶的水解活性。一种 Serpin 往往能抑制一种或几种丝氨酸蛋白酶。

使用抑制蛋白来抑制蛋白酶活性从能量学的角度似乎很不经济,因为每合成一个抑制蛋白分子需要消耗大量的 ATP。但由于蛋白酶的激活是不可逆的,当机体不再需要它们的时候,应该有一套快速将其灭活的方法,这对凝血系统和补体系统格外重要,而 Serpin 的存在正好能够满足这样的要求。正因为 Serpin 这种特性,所以它们在血浆中的含量极为丰富,最多可占到血浆总蛋白的 20%。倘若缺乏某种 Serpin,即可致病。

e12-1 丝氨酸蛋白酶与相应的抑制剂

在体内,嗜中性粒细胞(neutrophil)为了修复炎症的组织,经常向外分泌弹性蛋白酶。但如果弹性蛋白酶从炎症修复的地方扩散到肺泡,就会水解肺泡壁上的弹性蛋白(elastin)。为了防止这种情况的发生,肝细胞会分泌一种叫α1-抗胰蛋白酶的 Serpin。这种 Serpin 能与弹性蛋白酶结合,使其失活,来保护肺泡壁的完整。在α1-抗胰蛋白酶的分子上有一个关键的 Met 残基,它充当弹性蛋白酶的诱饵。一旦弹性蛋白酶结合上来,就会像其他丝氨酸蛋白酶遇到 Serpin 一样而失活。如果这个关键的 Met

残基因为基团突变而被其他氨基酸取代，则导致 Serpin 上的诱饵失效，于是弹性蛋白酶不再"上钩"，而是去不断地水解肺泡壁上的弹性蛋白，就会诱发肺气肿(emphysema)。吸烟可导致 α1- 抗胰蛋白酶分子上充当诱饵的 Met 残基发生氧化而失去活性，因此也可以诱发肺气肿。

Quiz7 到此为止，你学到了机体有哪几种方式可控制胰蛋白酶活性？

除了 Serpin 作为丝氨酸蛋白酶的抑制剂以外，细胞凋亡蛋白抑制剂(inhibitor of apoptosis protein，IAP)是胱天蛋白酶的抑制蛋白，金属蛋白酶组织抑制剂(tissue inhibitors of metalloproteinase，TIMP)是金属蛋白酶的抑制蛋白，半胱氨酸蛋白酶抑制剂(Cystatin)是溶酶体组织蛋白酶的抑制蛋白。IAP 与胱天蛋白酶的结合会抑制后者的活性从而阻断细胞凋亡，TIMP 通过其中的一个结构域与 Zn^{2+} 的作用而抑制金属蛋白酶的活性，Cystatin 在细胞内的功能是保护细胞，防止从溶酶体"泄漏"出来的组织蛋白酶破坏细胞。

使用激活蛋白来调节酶活性的一个重要例子是周期蛋白(cyclin)激活与调节细胞周期有关的蛋白质激酶。受周期蛋白激活的蛋白质激酶统称为依赖于周期蛋白的激酶(cyclin-dependent kinase，CDK)。已在真核细胞内发现多种不同的 CDK 和不同的周期蛋白，并且不同的 CDK 一般受不同的周期蛋白激活。在一个细胞周期中，各种 CDK 浓度的变化不大，但各自的周期蛋白搭档的变化很大。因此，在一个细胞周期的不同阶段，出现的周期蛋白不一样，而不同的周期蛋白激活不同的 CDK。正是由于不同的 CDK 按照一定的次序先后激活，才导致细胞周期从能一个阶段过渡到另外一个阶段。另外，真核细胞内广泛存在的钙调蛋白在与 Ca^{2+} 结合以后，能够作为许多酶的激活蛋白(参看第十七章"激素及其受体介导的信号转导")。

e12-2 各种 CDK 与其周期蛋白搭档

Quiz8 生的豆浆最好不要喝，这是为什么？

五、聚合与解离

某些酶具有单体和聚合体两种形式，但只有一种形式有活性。因而调节两种形式的互变也就成了调节酶活性的一种手段。

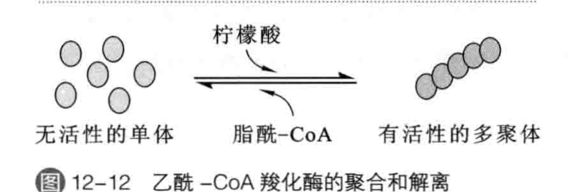


图 12-12 乙酰 -CoA 羧化酶的聚合和解离

通过这种方式调节酶活性的一个典型例子是真核细胞参与脂肪酸合成的乙酰 -CoA 羧化酶，该酶有单体和多聚体两种形式，单体无活性，多聚体才有活性。在细胞内，柠檬酸促进多聚体的形成，而脂酰 -CoA 则促进多聚体解离成单体(图 12-12)。

另外一个例子是哺乳动物脑垂体细胞内的丙酮酸激酶，该酶具有单体和四聚体两种形式，其中四聚体有酶活性，而单体具有结合甲状腺素的活性。这两种形式的转变由 1,6- 二磷酸果糖控制。1,6- 二磷酸果糖作为这种酶的前馈激活剂，当其水平上升的时候能促进具有酶活性的四聚体的形成；相反，如果 1,6- 二磷酸果糖水平下降，则有利于无酶活性但能够结合甲状腺素的单体形成。

框 12-1 生化趣事——"分子捕鼠器"的威力

Serpin 实际上是一种天然的自杀型抑制蛋白，它通过诱导酶的"自杀"来不可逆地抑制丝氨酸蛋白酶的活性。所有的 Serpin 都具有一个三股 β 折叠和 8~9 段 α 螺旋(从 N 端到 C 端的方向依次编号为 A~I)。分子表面具有一段外露的环，称为反应中心环(reactive centre loop，RCL)，其内部包括了特异性决定区和底物区。

Serpin 作用的基本步骤是(图 12-13)：首先它通过 RCL 模拟蛋白酶正常底物的结构而与酶活性中心结合。靶蛋白酶与它结合以后，将其视为底物，但只发生第一个半反应，即酶活性中心的 Ser 残基的羟基对 RCL 内的一个肽键做亲核进攻，形成脂酰化酶共价复合物，并释放位于肽键右端的肽段。正常的丝氨酸蛋白酶的底物随后会发生第二个半反应，促使酶与底物之间形成的共价键断裂。但是，Serpin 的 RCL 被切割以后立刻发生剧烈的构象变化。RCL 从分子的顶部，插入到 β 折叠之中，成为其中的第四个 β 股。下移的 RCL 迫使蛋白酶也从 Serpin 的顶部移至底部。蛋白酶的构象也因此发生变化，进而使得脂酰化

的酶难以水解，无法发生第二个半反应。于是蛋白酶始终与 Serpin 共价相连，两者“同归于尽”。有人把 Serpin 比作是一种分子捕鼠器（molecular mousetrap），丝氨酸蛋白酶比作是老鼠，RCL 相当于是捕鼠器上的弹簧，其内的一个甲硫氨酸残基的侧链相当于诱饵。当老鼠受到诱饵的诱惑，禁不住咬上诱饵的时候，弹簧在瞬间滑动到另外一处，将老鼠牢牢夹住，直至其最后死亡。不过与真正的捕鼠器不同的是，Serpin 的使用是一次性的。

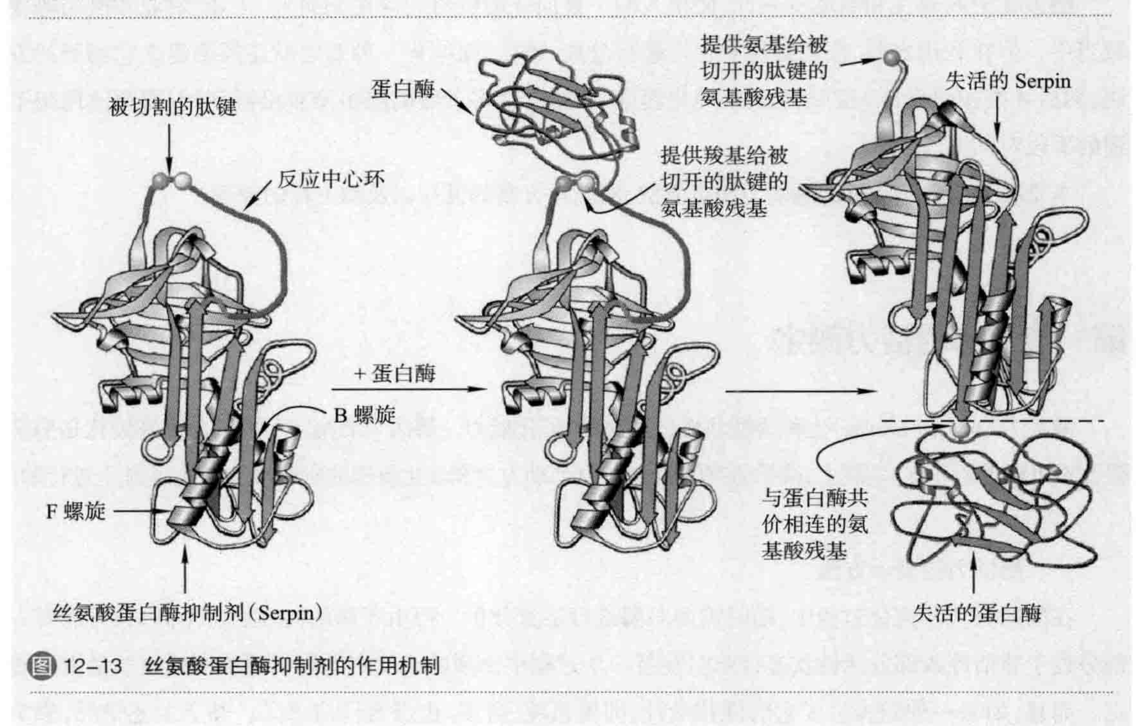


图 12-13 丝氨酸蛋白酶抑制剂的作用机制

e12-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Allosteric_regulation（维基百科有关别构调节的内容）
2. http://web.virginia.edu/Heidi/chapter15/chp15.htm（美国弗吉尼亚大学一个网站，有关酶的特异性和调节机制的内容）
3. http://www.pearsonhighered.com/mathews/ch11/c11a.htm（Mathews 生物化学网络课程有关酶活性调节的内容）
4. http://www.biochem.arizona.edu/classes/bioc462/462a/NOTES/ENZYMES/enzyme_regulation.html（美国亚利桑那大学一个生物化学课程网站，显示的 Stryer 编著的《生物化学》有关酶活性调节的内容）

参考文献：

1. Bu Z, Callaway D J. Proteins MOVE! Protein dynamics and long-range allostery in cell signaling. *Adv in Protein Chemistry and Structural Biology*, 83: 163–221, 2011.
2. Satyanarayana A, Kaldis D. Mammalian cell-cycle regulation: several Cdks, numerous cyclins, and diverse compensatory mechanisms. *Oncogene*, 28: 2 925–2 939, 2009.
3. Barford D. Structural studies of reversible protein phosphorylation and protein phosphatases. *Biochem Soc Trans*, 27: 751–766, 1999.
4. Kantrowitz E R, Lipscomb W N. *E.coli* aspartate transcabamoylase: The relation between structure and function. *Science*, 241: 669–674, 1998.
5. Neurth H. The versatility of proteolytic enzymes. *J Cell Biochem*, 32: 35–49, 1986.
6. Monod J, et al. On the nature of allosteric transitions: A plausible model. *J Mol Biol*, 12: 88–118, 1965.

第十三章　酶的应用及研究方法

酶催化的高效性和高度专一性，使得人们一直向往能够将它应用到研究、工业、农业和医药等领域当中。但在利用之前，首先必须对它们进行分离、纯化，而纯化一种酶之前还需要建立它的测活方法，然后才能按照设计好的方案进入纯化程序。此外，许多天然的酶很难直接被利用，需要使用酶工程的手段对其进行改造。

本章将主要介绍常见的酶活力测定方法、酶纯化方案的设计以及酶工程的原理。

第一节　酶的活力测定

酶活力(enzyme activity)也称为酶活性，是指酶的催化能力。酶活力的测定与酶的分离和纯化是酶学研究不可缺少的环节。实际上，酶学研究的诸多方面(如动力学和催化机理的研究)都是在此基础上进行的。

一、酶活力的表示方法

在酶的分离和纯化过程中，随时需要对酶进行定量分析。但由于酶的纯度通常不高，且可能有一部分处于非活性或部分活性状态，因此，在某一个过程中出现的或使用的酶量很难用绝对的量纲去确定。而且，如果一种酶丧失了它的催化活性，即使再纯、再多，也没有任何意义。基于上述情况，酶学家在对酶进行定量的时候，通常需要将它与酶的活力联系起来。他们经常使用的是活力单位(activity unit，U)。1964 年，国际生物化学联合会(the International Union of Biochemistry，IUB)采纳了国际酶学委员会于 1961 年给酶活力单位下的定义：1 个 U 是指在最适条件下每分钟催化 1 μmol 底物转化的酶量。一般而言，这样的 1 个单位相当于 10^{-11}~10^{-6} kg 的纯酶或 10^{-7}~10^{-4} kg 的工业酶制剂。除了上述对酶活力单位的定义以外，国际纯粹与应用化学联合会(the international union of pure and applied chemistry，IUPAC)在 1972 年还推荐了另外一种叫 katal(kat)的单位，它被定义为每秒钟催化 1 分子底物转化的酶量($1\ kat=6\times10^7\ U$，$1\ U=16.67\times10^{-9}\ kat$)。这样的单位实用性仍然不强，所以至今也没有被广泛采纳，很多研究者还是按照各自的需要来定义一种酶的活力单位。

上述两种表示酶活力的方法都不能反映一个酶制剂的纯度，一个酶活力高的样品可能纯度并不高，相反，一个酶活力低的样品也可能纯度很高。为了能够更好地说明一种酶的纯度，IUPAC 推荐使用比活性或比活力(specific activity)来表示，它是指单位质量(通常是每毫克)酶所含有的活力单位数。显然，对同一种酶而言，比活力越高，酶纯度就越高。正因为如此，在进行酶纯化的时候，需要时刻关注比活性的变化。当一种酶的比活性不能再增加的时候，此酶可视为高纯度。然而，还需要注意的是，酶的比活力与酶的稳定性有密切的关系。任何一种酶的比活性都会随着时间的推移而下降，稳定性越差，比活性下降得就越快。

二、酶活力测定的方法

测定一种酶的活力实际上就是测定它所催化的化学反应的最佳反应速率。而测定反应速率的方法原则上有两种，一种是检测单位时间内底物的减少量，另一种是测定单位时间内产物的增加量。使用后一种方法更为常见，原因是当测定反应的初速率时，产物量的变化是从无到有，其变化更为敏感。

既然酶促反应速率受到多种因素的影响，那么在检测酶活力的时候，就应该尽可能让酶本身的催化能力充分地展示出来，一切不利于酶促反应的因素都应当被降到最低。总的原则是①反应条件为

最适条件,包括最适 pH、最适温度和最适离子强度等;②反应速率为初速率;③底物浓度过量。

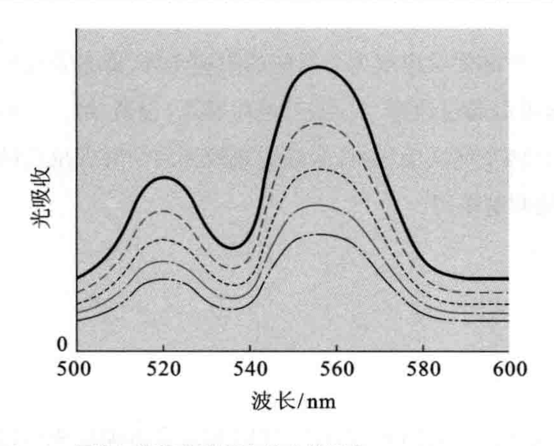
在采用上述原则以后,测定一种酶活力最为关键的一步就是选择一种检测底物量或产物量变化的方法。然而,不同酶促反应的底物或产物在结构上千差万别,有的很容易检测,有的则不然。于是,人们根据底物或产物的结构和性质设计了不同的检测方法,综合起来,通常有以下 3 种。

(一) 直接测定法

使用直接(direct assay)测定法的酶促反应的产物或底物会产生明确的可检测信号,只需利用专门的仪器就可以直接测定出它们的变化。例如使用专门的放射性同位素计数仪,可测定被放射性同位素标记的底物量的变化;使用分光光度计(spectrophotometry),可在特定的波长检测具有光吸收的底物或产物量的变化;使用荧光光谱测定法(spectrofluorimetry),可测定能发出荧光的底物或产物量的变化。

以细胞色素氧化酶为例(详见第二十章“生物氧化”),它催化的反应是:细胞色素 c(Fe^{2+})→细胞色素 c(Fe^{3+})。由于底物(还原型细胞色素 c)和产物(氧化型细胞色素 c)具有不同的光吸收性质,前者在 560 nm 有吸收峰,后者无,因此通过分光光度计测定在单位时间内光吸收在 560 nm 处的减少,就可以测定出这种酶的活力(图 13-1)。

Quiz1 如何直接测定乙醇脱氢酶、逆转录酶和蛋白质激酶的活性?



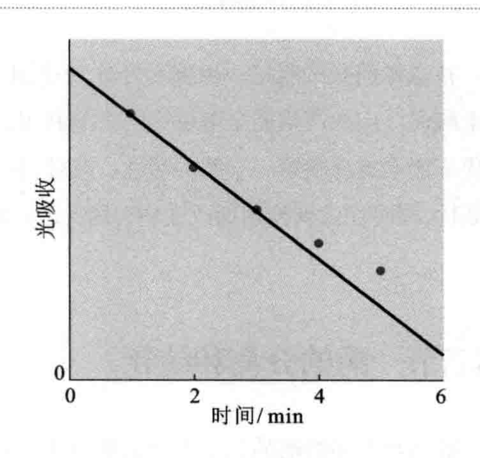


图 13-1 细胞色素氧化酶活力的直接测定

图 13-2 二氢乳清酸脱氢酶的间接测定

(二) 间接测定法

使用间接测定法(indirect assay)测定的酶促反应的产物和底物都无法提供明确的可检测信号,因此不能直接测定,但可以将释放的产物与一个能产生特定的可检测信号的非酶促反应偶联在一起,进行间接的测定。以二氢乳清酸脱氢酶(dihydroorate dehydrogenase)测定为例(图 13-2),反应能以二氢乳清酸和氧化态辅酶 Q 为底物,产生乳清酸和还原态辅酶 Q。如果反应系统中存在氧化态的二氯酚靛酚(dichlorophenol indophenol,DCPIP),它会与还原态辅酶 Q 起反应,将辅酶 Q 重新氧化成氧化态,而它自身则还原成还原态。由于还原态的 DCPIP 无色,氧化态的 DCPIP 呈深蓝色,因此可以利用 DCPIP 颜色的变化间接测定二氢乳清酸脱氢酶的活性。

(三) 偶联测定法

偶联测定法(coupled assay)的原理是将一个难以测定的酶促反应跟一个容易测定的酶促反应

偶联在一起，通过第二个酶反应测定第一个酶的活性。以己糖激酶催化的反应为例，该反应的产物 6-磷酸葡糖可作为第二个 6- 磷酸葡糖脱氢酶的底物。在此酶的催化下，6- 磷酸葡糖脱氢转变成 6- 磷酸葡糖酸内酯，同时 $NADP^+$ 被还原成 NADPH。由于 NADPH 在 340 nm 有吸收峰，因此可以通过分光光度计测定 340 nm 光吸收的增加来间接测定己糖激酶的活性（图 13-3）。

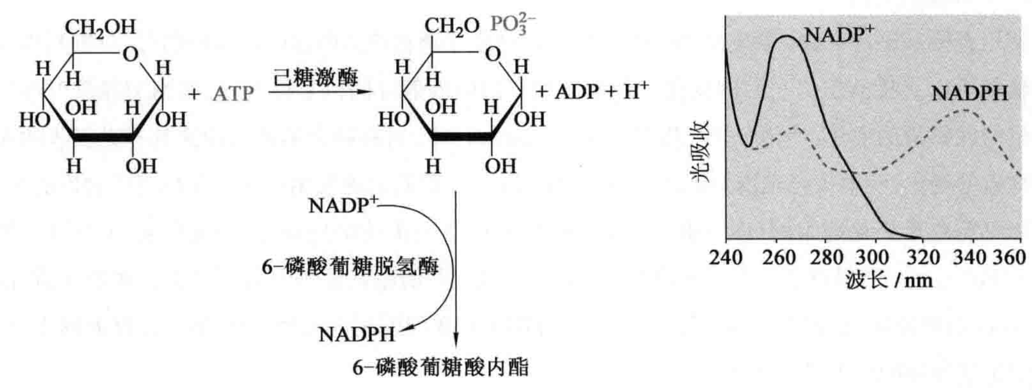


图 13-3　己糖激酶活力的偶联测定

在选择偶联法测定一种酶活力时，必须让第一个酶促反应成为总反应的限速步骤，即必须保证第一个酶促反应的产物能立即被第二个酶利用。这可以通过让第二个酶过量而实现，否则，第二个酶促反应不能准确反映第一个酶的活性。此外，在设计两个偶联反应时，最好能兼顾到两个酶促反应的最适条件，同时注意到抑制剂或底物对两个酶可能的抑制作用。

Quiz2　如何使用偶联法来测定顺乌头酸酶的活性？

第二节　酶的分离和纯化

既然绝大多数酶的化学本质是蛋白质，那么纯化蛋白质的各种方法、策略和注意事项完全可以用于酶的分离和纯化。

图 13-4 为一种酶典型的分离纯化流程图，* 表示可以选择的手段，它们的前后次序需要根据具体酶的性质做适当调整。需要注意的是，每一步完成以后取得的进展可以通过以下几种方法进行鉴定：

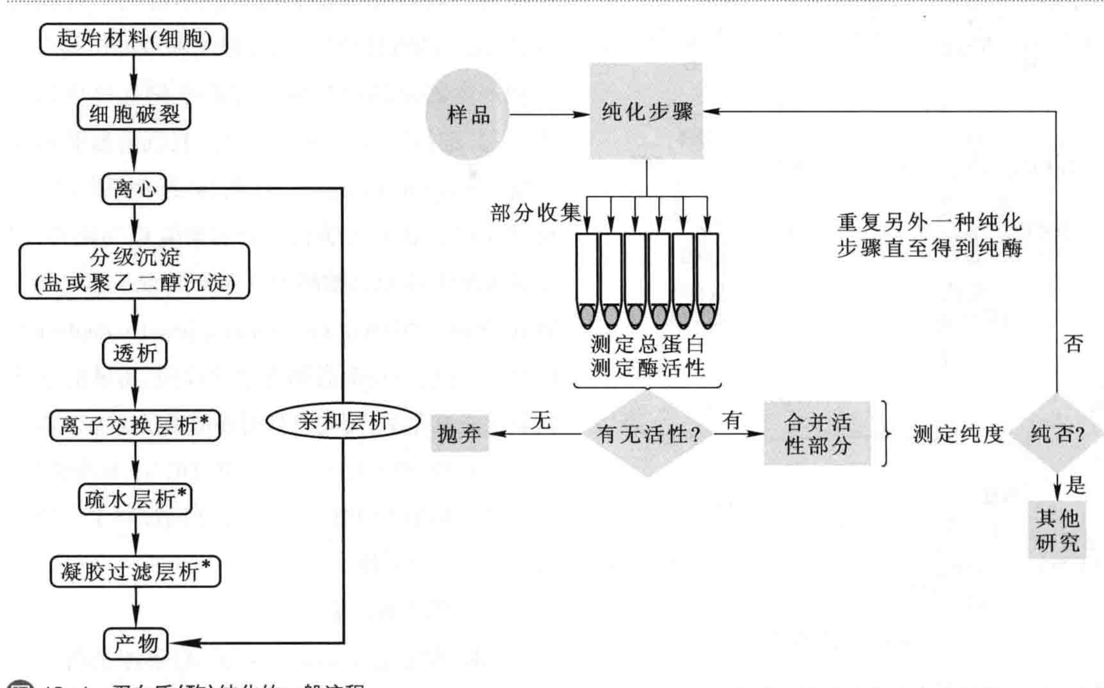


图 13-4　蛋白质（酶）纯化的一般流程

①测定回收到的总蛋白量;②测定回收到的酶总活性;③凝胶电泳和比活性的测定,以观测目标酶纯度和杂蛋白条带的变化。

当纯化完成以后,不要忘记绘制一张酶纯化表(表 13-1),在表中需要注明每一步纯化得到的数据:①酶溶液体积(ml);②酶溶液蛋白质含量(mg/ml);③酶溶液活性(U/ml);④酶总量或酶总活性(U)= 酶活力(U/ml)× 体积(ml);⑤比活性(U/mg)= 酶活力(U/ml)/ 蛋白质量(mg/ml);⑥总蛋白(mg)= 酶溶液蛋白质含量(mg/ml)× 体积(ml);⑦得率(%)= 每一步纯化后的酶总量 / 每一步纯化之前的酶总量 ×100%;(8)纯化倍数 = 每一步纯化后的酶比活性 / 每一步纯化之前的酶比活性。

Quiz3 你如何用实验的方法确定一种酶有没有四级结构?

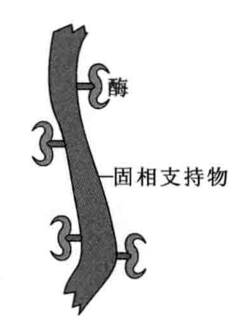
表 13-1 酶纯化表

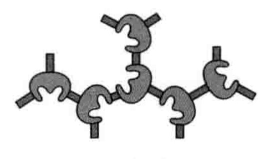
纯化步骤	酶溶液体积 /mL	总蛋白 /mg	总活性 /U	比活性 /(U·ml^{-1})	纯化倍数	得率 /%
粗细胞抽取物	1 400	10 000	100 000	10	1	100
凝胶过滤	90	400	80 000	200	20	80
离子交换	80	100	60 000	600	3	75

第三节 酶工程

许多天然酶本身存在一些不尽如人意的性质,如稳定性差、抗原性强、副作用大、含量低和反应条件特殊等。相对于科研和生产的要求,酶的这些缺陷大大地限制了它们的应用,因此,研究人员一直在想方设法利用各种手段,按照自己的意愿去改造它们,甚至在对已有的酶结构和功能认识的基础上,去创造自然界根本不存在的酶,以造福人类。然而,目前对生物大分子的了解(特别是对酶)远远落后于对小分子的了解,其主要原因是由于酶的性质和功能由其精确的三维构象决定,而现在对于指导蛋白质折叠的详细机制以及蛋白质结构和功能之间的关系的理解还很不完全。这就需要人们进行更多的基础研究,同时需要多种学科,特别是化学、免疫学、结构生物学、生物信息学、分子生物学和遗传学等在此领域的交叉和融合。

就目前的水平,从头设计一种酶还不切实际,但对已经存在的酶进行各种形式的修饰和改造却是切实可行的。修饰和改造的手段可以是化学的,也可以是生物学的,前者为化学酶工程,后者为生物酶工程。按照修饰或改造的具体手段,酶工程可以分为固定化酶(immobilized enzymes)、人工酶(artificial enzymes)、定点突变酶、杂交酶和抗体酶等。





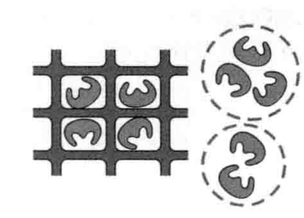


图 13-5 酶固定化的三种方法

一、固定化酶

酶的商业化应用越来越普及,尤其是在洗涤、食品和酿造工业上。但是,如果直接使用可溶性的酶制剂往往存在以下问题:①溶解在反应系统中,难以回收;②某些产物能抑制酶活力(反馈抑制);③不稳定。

固定化酶就是针对上述问题而发展起来的一项技术,其原理是指将一种可溶性酶与不溶性的有机或无机基质结合,或者将其包埋到特殊的具有选择透过性的膜内,从而提高酶的稳定性,也便于重复和持续使用。与可溶性酶相比,固定化酶具有方便、经济和稳定等优点,很好地解决了直接使用而产生的一些问题。

酶的固定化方法有载体结合(carrier-binding)、交联(cross-linking)和包埋(entrapping)三种(图 13-5)。

(一)载体结合

载体结合的方法是最早的酶固定化技术。在这种方法中,与载体结合的酶量和固定以后的酶活力取决于载体的性质。载体的选择又取决于酶的性质以及颗粒大小、比表面积、亲水基团与疏水基团

的摩尔比和化学组成等因素。

用来固定酶的载体最常见的有多糖衍生物(如纤维素、葡聚糖和琼脂糖)和聚丙烯酰胺凝胶。根据酶的结合方式,载体结合还可以进一步分为物理吸附(physical adsorption)、离子结合(ionic binding)和共价结合(covalent binding)。

(二)交联

交联是通过双功能或多功能试剂的作用,在酶分子之间或酶与不溶性支持物分子上的功能基团之间形成共价键的一种酶固定方法,通常会与其他方法结合起来使用。如果交联在酶分子之间进行,则代价昂贵,效率低下,原因是一些酶分子将不可避免地起支持物的作用。

使用最多的双功能试剂是戊二醛(glutaraldehyde)。由于交联反应一般在较为激烈的条件下进行,而这样的条件会改变酶活力中心的构象,因此,有时会造成酶活力的显著下降。

(三)包埋

包埋是将酶分子整合到半透性胶形成的网格内或用半透性的多聚物膜包被酶的一种固定化方法。在包埋中使用的胶或膜所起的作用都是截留酶,同时允许底物和产物自由通过。

与共价结合和交联不一样的是酶分子在包埋中并不与胶或膜结合,这样的性质也使它的应用更为广泛。但在包埋的时候,化学聚合的反应比较剧烈,会导致酶活力的损失。因此,各种酶在包埋中选择好最合适的条件是非常必要的。

根据包埋的方式,包埋又可以分为网格型(lattice-type)和微胶囊型(microcapsule-type)。前者将酶包在交联的水不溶性多聚物的间隙内,后者则将酶包在半透性的多聚物膜内。

二、人工酶

人工酶也称为人工合成酶(synzyme),它们一般是具有类似酶活力的合成多聚物或寡聚物,有时还包括具有酶活力的天然蛋白的衍生物(并非固定化酶)。1977年Dhar报道,人工合成的八肽序列Glu-Phe-Ala-Glu-Glu-Ala-Ser-Phe具有溶菌酶的活性,其活力为天然酶的一半。1990年,Steward等人使用胰凝乳蛋白酶的底物酪氨酸乙酯作为模板,用计算机模拟胰凝乳蛋白酶的活性位点,构建出一种由73个氨基酸残基组成的多肽,其活性部位由Ser、Asp和His组成。此肽对烷基酯底物的活力为天然胰凝乳蛋白酶的1%,同时还显示了底物特异性以及对胰凝乳蛋白酶抑制剂的敏感性等。还有人将$(Ru(NH_3)_5)^{3+}$连到肌红蛋白表面的3个His残基上,可以将其从氧气的载体转化成一种氧化酶——氧化抗坏血酸同时还原分子氧,效果与天然的抗坏血酸氧化酶(ascorbate oxidase)相差不多。

Quiz4 为什么人工酶一般遵守米氏饱和动力学?

人工酶必须具备两个结构要素,一个是底物结合位点,另一个是催化位点。一般情况下,得到底物结合位点相对简单,而得到催化位点比较困难。在设计两个位点的时候,可分别考虑,但要是底物结合位点能很好地结合反应过渡态类似物,那么这样的位点往往也具有催化活性。一般而言,人工酶遵守米氏反应动力学。

许多人工酶是通过模拟天然酶与底物的结合和催化过程而得到的,这些酶也称为模拟酶。例如,对一些天然或人工合成的化合物引入某些活性基团,使其具有酶的行为。目前用于构建模拟酶的这类酶模型分子有环糊精、冠醚、穴醚、笼醚和卟啉等。利用环糊精已成功地模拟了胰凝乳蛋白酶、核糖核酸酶、转氨酶、碳酸酐酶等。1985年,Bender等人利用β-环糊精的空穴作为底物的结合部位,以连在环糊精侧链上的羧基、咪唑基及环糊精自身的一个羟基共同构成催化中心,构建成了胰凝乳蛋白酶的模拟酶。

三、定点突变酶

定点突变是利用重组DNA技术,在基因水平上对编码酶的核苷酸序列进行定点突变,以使酶在特定位置的氨基酸序列发生变化,再经过筛选从而得到“新酶”的过程(详见第四十四章“重组DNA技术”)。

通过对酶基因的定点突变可以改变酶的性质(如酶活力、稳定性、底物专一性和对辅酶的依赖性等),从而得到具有新性状的酶。例如,将枯草杆菌蛋白酶的 Asp99 和 Glu156 替换成 Lys 后,使这个酶在 pH 7 和 pH 6 时的活力分别提高了 1 倍和 10 倍。

四、杂交酶

现代分子生物学技术的发展已允许人们将两种不同的生物分子融合在一起,以获得具有新性质、新功能的杂合分子。杂交酶就是酶与其他生物分子融合在一起的产物。以图 13–6 为例,将一段特定序列的寡聚核苷酸"嫁接"到葡萄球菌核酸酶(staphylococcal nuclease)的 Cys116 残基上,形成的是一种蛋白质 / 核酸杂交酶。由于接到酶分子上的碱基序列能与细胞内的 DNA 分子上特定的碱基序列互补,因此,这样的杂交酶能通过碱基序列的互补对一种 DNA 分子实行定点切割。显然,天然的葡萄球菌核酸酶无此特异性。

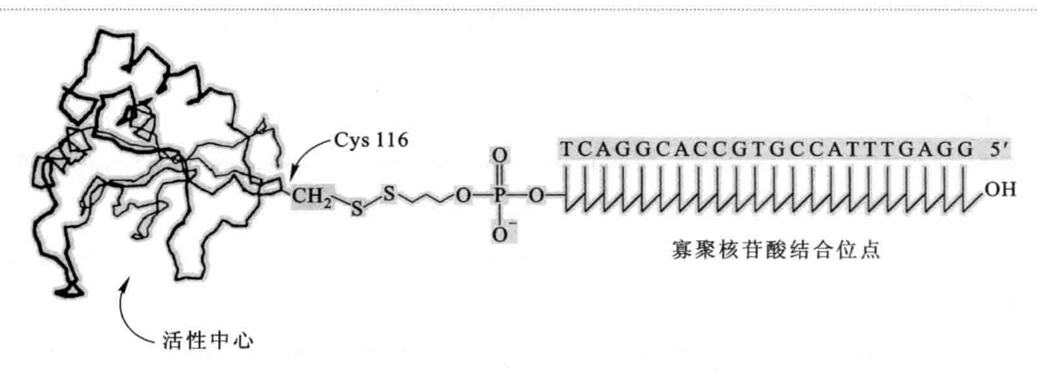


图 13–6 衍生于葡萄球菌核酸酶的蛋白质 / 核酸杂交酶

五、抗体酶

抗体酶也就是催化性抗体,它是抗体的高度特异性与酶的高效催化性巧妙结合的产物,其本质上是一类在可变区赋予了酶活力的免疫球蛋白。

迄今为止,抗体酶所催化的反应除了酯、羧酸和酰胺的水解外,还有酰胺形成、氧化还原、光诱导裂解和聚合、酯交换等。这些抗体酶催化反应的专一性相当于甚至超过天然的酶,催化速率少量可达到酶催化的水平,但多数催化能力不是很高。

框 13–1 生化动态——抗体酶的新来源

现在除了使用人工设计的过渡态类似物作为半抗原,然后用它去免疫动物获得抗体酶以外,人们已发现另外一种获得抗体酶的方法,就是从患自身免疫性疾病患者的血清中去寻找天然的抗体酶。

这种获得抗体酶的想法源自 Paul 等人的发现。在 1989 年,Paul 发现一些哮喘病患者血清中的 IgG,可以特异性催化一种叫血管活性肠肽(vasoactive intestinal peptide,VIP)的激素水解。这种抗体酶对 VIP 的亲和力非常高,两者的解离常数 K_d 约为 0.76 nmol/L。VIP 上位于 14~22 号位的肽键对这种抗体酶敏感。进一步研究还表明,这种抗体的轻链单独就能水解 VIP。对这种抗体酶的催化机制的研究表明,它实际上就是一种丝氨酸蛋白酶。

显然,这样的抗体比仅仅是与抗原可逆结合的抗体中和抗原的能力要强得多。后来有人在淋巴瘤性甲状腺肿患者和系统性红斑狼疮患者的血清中,同时发现可水解甲状腺球蛋白的抗体酶。还有人在多髓瘤患者的血清中发现具有凝血酶原活性的抗体轻链。

除了针对蛋白质的抗体酶以外,一些可水解 DNA 和 RNA 的天然抗体酶也相继被发现。最神奇的是,有人在健康产妇的母乳中发现具有激酶活性的抗体,这类抗体酶可催化母乳中的酪蛋白发生磷酸化。

e13–1 化学修饰酶

Quiz5 你认为哪些氨基酸的侧链基团可以用来进行共价修饰?

原因在哪里？主要是因为抗原－抗体间的结合是相当肤浅的接触。虽然这样已经足够作为免疫识别之用，但在酶催化上却有所不足。事实上，酶的催化反应比单纯的抗原－抗体结合要复杂得多，牵涉到许多电子的转移与空间的扭动。因此，酶活力中心多深埋在分子内部，多数为一个口袋式的结合区，以避开环境中水分子的干扰，使得底物或过渡状态能与活性区有更密切的结合，而不只是表面的接触而已。

除了以上几种酶工程手段以外，化学修饰酶也是一种常见的形式。这种酶工程是利用化学的手段，对酶分子上的氨基酸侧链基团进行修饰，以改善酶的性能。

e13-2 本章小结填填看

推荐网址：

1. http://wolfson.huji.ac.il/purification/Purification_Protocols.html（内有各种生物分子纯化的方法）
2. http://en.wikipedia.org/wiki/Immobilized_enzyme（维基百科有关固定化酶的内容）

参考文献：

1. Takahashi N, et al. In vitro abzyme evolution to optimize antibody recognition for catalysis. *Nature Biotechnology*, 19: 563–567, 2001.
2. Zhou G W, et al. Crystal structure of a catalytic antibody with a serine protease active site. *Science*, 265: 1 059–1 064, 1994.
3. Zhong Z, et al. Active-site-directed modification of subtilisin *J Am Chem Soc*, 113 (6): 2 259–2 263, 1991.

第十四章 维生素与辅酶

维生素(vitamin)是维持生物体正常生命活动必不可少的一类小分子有机化合物。虽然机体对它们的需要量甚少(一个人每日需要量在微克或毫克级),但由于它们不能在体内合成,或者虽能合成但合成的量难以满足机体的需要,必须通过饮食等手段获取。虽然它们在体内既不是构成细胞组织的原料,也不是供能的物质,但是,它们在代谢调节、促进生长发育和维持生理功能等方面却发挥着十分重要的作用。因此,人体如果长期缺乏某种维生素,就会出现相应的维生素缺乏症。

维生素的种类多、来源广、功能多样,其化学结构差别也很大。为方便起见,通常按溶解性质将其分为脂溶性维生素(fat-soluble vitamin)和水溶性维生素(water-soluble vitamin)两大类。这两类维生素的主要差别参看表 14-1。

e14-1 各种各样的维生素

表 14-1 脂溶性维生素与水溶性维生素的比较

类别	脂溶性维生素	水溶性维生素
溶解性质	不溶于水,溶于有机溶剂	溶于水
吸收	被小肠吸收后,先进入淋巴循环,然后再到血液	被肠道吸收后直接进入血液
血液运输	需要载体蛋白的帮助	游离的形式
贮存	量多时与脂肪贮存在一起,难以排泄	量多时经肾排泄出去
毒性	大量服用时容易达到毒性水平	难以达到毒性水平
实例	维生素 D、A、K 和 E	B 族维生素和维生素 C

本章将主要介绍各种维生素的结构和生理功能,并重点讨论它们与辅酶或辅基的关系。

Quiz1 你如何提高机体对脂溶性维生素的吸收效率?

第一节 水溶性维生素

水溶性维生素包括 B 族维生素和维生素 C,它们在生物体内能够直接作为辅酶或辅基,或者转变为辅酶或辅基,参与物质代谢和能量代谢。当水溶性维生素缺乏时,机体的代谢会出现障碍,最容易受到影响的是生长和分裂旺盛的细胞和组织,如上皮细胞和血细胞。不同的水溶性维生素的缺乏往往会有一些交叉的症状,如皮炎(dermatitis)、舌炎(glossitis)、口角炎(cheilitis)和腹泻(diarrhea)。由于神经组织的活动非常依赖于持续的能量供应,尤其是来自糖氧化分解所释放出的能量,因此,在很多情况下,缺乏水溶性维生素也会影响到神经系统的功能,主要症状有外周神经炎(peripheral neuropathy)、忧郁(depression)、精神错乱(mental confusion)和运动失调等。

一、B 族维生素

这是一个大家族,包括维生素 B_1、维生素 B_2、维生素 PP、维生素 B_6、泛酸、生物素、叶酸、硫辛酸和维生素 B_{12}。它们的共同特点有:①在自然界经常共同存在,最丰富的来源是酵母、蔬菜和动物肝脏;②从低等的微生物到高等动物和人类都需要它们作为营养要素;③在生物体内主要作为辅酶或辅基参与代谢;④从化学结构上看,大都含有 N;⑤从性质上看,此类维生素易溶于水,对酸稳定,易被碱或热破坏。

(一) 维生素 B_1

维生素 B_1 是第一个被发现的维生素,其化学结构如图 14-1 所示,因具有含 S 的噻唑环和含氨基的嘧啶环故又名为硫胺素(thiamine)。

维生素 B_1 易被小肠吸收，在细胞内受激酶的催化被磷酸化成焦磷酸硫胺素（TPP）。TPP 是体内催化 α- 酮酸氧化脱羧的酶的辅酶，也是磷酸戊糖途径中转酮酶的辅酶（分别参看第二十三章“三羧酸循环”和第二十四章“磷酸戊糖途径”），在反应中直接作为亲核试剂参与催化。当维生素 B_1 缺乏时，由于 TPP 合成不足，丙酮酸和 α- 酮戊二酸的氧化脱羧以及磷酸戊糖途径即发生障碍，这必然会导致糖的氧化利用受阻。在正常情况下，神经组织的能量供应依赖于糖的氧化分解，因此维生素 B_1 缺乏首先会影响神经组织的能量供应，并伴有丙酮酸及乳酸等在神经组织中的堆积，出现手足麻木、四肢无力等多发性外周神经炎的症状。重者引起心跳加快、心脏扩大和心力衰竭，临床上称为脚气病（beriberi），因此维生素 B_1 又被称为抗脚气病维生素。

硫胺素

焦磷酸硫胺素（TPP）

图 14-1　维生素 B_1 及其的辅酶（TPP）的化学结构

维生素 B_1 还有抑制乙酰胆碱酯酶的作用。乙酰胆碱酯酶能催化神经递质——乙酰胆碱的水解。因此，当缺乏维生素 B_1 时，乙酰胆碱酯酶活性将增强，乙酰胆碱水解加速，神经传导会受到影响，可造成胃肠蠕动缓慢、消化液分泌减少、食欲缺乏和消化不良等症状。反之，补充维生素 B_1 则可增加食欲、促进消化。富含 B_1 的食品有肉类、绿叶素菜、谷物及麦片等。

（二）维生素 B_2

维生素 B_2 是由核糖醇（ribitol）与 7,8- 二甲基异咯嗪（iso-alloxazine）结合而成（图 14-2）。氧化型的维生素 B_2 呈黄色，因而又名为核黄素（riboflavin）。异咯嗪环上的 N1 和 N5 可加氢和脱氢，具有可逆的氧化还原特性，而这一特点与核黄素的主要生理功能直接相关。

核黄素在体内经磷酸化作用，可转变为黄素单核苷酸（flavin mononucleotide，FMN）和黄素腺嘌呤二核苷酸（flavin adenine dinucleotide，FAD）（图 14-2），它们构成各种黄酶或黄素蛋白的辅基参与体内生物氧化，作用机理详见第二十章“生物氧化”。几种重要的黄酶包括：NADH 脱氢酶（参看第

7,8-二甲基异咯嗪

核糖醇

核黄素

FMN

FAD

图 14-2　维生素 B_2 及其衍生的辅基（FMN 和 FAD）的化学结构

二十章“生物氧化”)、二氢硫辛酸脱氢酶及琥珀酸脱氢酶(参看第二十三章“三羧酸循环”)、脂酰辅酶A脱氢酶(参看第二十九章“脂肪酸代谢”)、3-磷酸甘油脱氢酶(参看第二十八章“脂肪、磷脂和糖脂的代谢”)、氨基酸氧化酶(参看第三十一章“氨基酸代谢”)和黄嘌呤氧化酶(参看第三十二章“核苷酸代谢”)。

维生素 B_2 缺乏时,主要症状为口角炎、舌炎、阴囊炎、皮疹及角膜血管增生和巩膜充血等。婴、幼儿缺乏它则会生长迟缓。富含 B_2 的食品有牛奶和乳制品、肉类、绿叶素菜、谷物及麦片等。

(三) 维生素 PP

维生素 PP 即维生素 B_3,包括尼克酸(烟酸,nicotinic acid)和尼克酰胺(烟酰胺,nicotinamide),两者均为吡啶衍生物,在体内可以相互转变。动物组织中维生素 PP 大多以尼克酰胺的形式存在。尼克酰胺是构成辅酶Ⅰ(NAD^+)和辅酶Ⅱ($NADP^+$)的成分之一(图 14-3),与 FMN 和 FAD 一样,也具有可逆的加氢和脱氢的特性,因此这两种辅酶也参与生物氧化,作为细胞内很多重要的脱氢酶的辅酶,例如,3-磷酸甘油醛脱氢酶和 6-磷酸葡糖脱氢酶分别以辅酶Ⅰ和辅酶Ⅱ作为辅酶(参看第二十二章“糖酵解”和第二十四章“磷酸戊糖途径”)。

图 14-3 维生素 PP 及其衍生的辅酶(NAD^+ 和 $NADP^+$)的化学结构

图 14-4 维生素 B_6 及其衍生的辅酶的化学结构

NAD^+ 除了可以作为许多脱氢酶的辅酶以外,在体内至少还能作为其他 3 种酶的底物参与相关的反应:①许多细菌(如大肠杆菌)的 DNA 连接酶(参看第三十三章“DNA 复制”);②真核细胞内依赖于 NAD^+ 的组蛋白去乙酰酶(参看第四十三章“真核生物的基因表达调控”);③具有 ADP-核糖基转移酶活性的霍乱毒素、百日咳毒素和白喉毒素(参看第十七章“激素及其受体介导的信号转导”)。

尼克酸在人体内可从色氨酸代谢产生,但色氨酸转变成尼克酸的量有限,不能满足机体的需要,因此仍需要从食物中获取。一般营养条件下,很少会出现缺乏维生素 PP 的情况。维生素 PP 缺乏时,主要表现为癞皮病(pellagra)。由于补充维生素 PP 可预防和治愈癞皮病,因此维生素 PP 又被称为抗癞皮病因子或抗癞皮病维生素。

(四) 维生素 B_6

维生素 B_6 包括吡哆醇(pyridoxine)、吡哆醛(pyridoxal)和吡哆胺(pyrodixamine)3 种形式(图 14-4),在体内可以相互转变。

Quiz2 为什么长期单食玉米有可能发生维生素 PP 缺乏病?

细胞内的维生素 B_6 在激酶的催化下经磷酸化作用转变为相应的磷酸酯,其中作为辅酶的主要是

磷酸吡哆醛(PLP)和磷酸吡哆胺。它们在体内参与氨基酸的转氨、消旋、某些氨基酸的脱羧、半胱氨酸的脱巯基作用(参看第三十一章“氨基酸代谢”)和糖原的磷酸化,此外还参与血清素、去甲肾上腺素、鞘磷脂以及血红素的合成。

维生素 B_6 在动植物中分布极广,同时,人体内的肠道细菌也能够合成它,因此人类尚未发现单纯的维生素 B_6 缺乏病。动物缺乏维生素 B_6 可引发与癞皮病相似的皮炎。

(五) 泛酸

泛酸(pantothenic acid)即维生素 B_5,是由 α,γ- 二羟 -β,β- 二甲基丁酸与 β- 丙氨酸通过酰胺键缩合而成的酸性物质,因其广泛存在于动、植物组织中,故名泛酸或遍多酸。

泛酸在体内与巯基乙胺、焦磷酸及 3′-AMP 磷酸结合成为辅酶 A(coenzyme A,CoA)(图 14-5)。辅酶 A 最重要的活性基团为巯基,在脂代谢中以它的巯基作为脂酰基的载体,故辅酶 A 常用 CoA-SH 表示。

4-磷酸泛酰乙胺

$HS-CH_2-CH_2-NH-C(=O)-CH_2-CH_2-NH-C(=O)-CH(OH)-C(CH_3)_2-CH_2-O-P(=O)(O^-)-O-P(=O)(O^-)-O-CH_2$-(3′-磷酸核糖)-腺嘌呤($NH_2$)

β-巯基乙胺　泛酸　3′-磷酸-核糖　3′-磷酸-ADP

图 14-5　泛酸及辅酶 A 的化学结构

(六) 叶酸

叶酸(folic acid)是由蝶酸(pteroic acid)和谷氨酸缩合而成,因在植物绿叶中含量丰富而得名。食物中的叶酸主要是蝶酸与寡聚谷氨酸的缩合物,即寡聚谷氨酸叶酸(图 14-6)。在消化道内,这种形式的叶酸经结合酶(conjugase)的作用,转变为单谷氨酸叶酸后才能被吸收。

细胞吸收叶酸的过程是受受体介导的。叶酸的受体是位于细胞膜上的一种糖基磷脂酰肌醇锚定蛋白(参看第十六章“脂质与生物膜”)。当叶酸与膜上的受体结合以后,就发生受体介导的内吞,但叶酸与受体在胞内遇到偏酸的环境即发生解离。

叶酸在细胞内的辅酶形式为 5,6,7,8- 四氢叶酸(FH_4 或 THF),其作用是参与体内“一碳单位”的转移(图 14-7),充当甲基(methyl)、亚甲基(methylene)、甲酰基(formyl)、甲川基(methenyl)和亚胺甲基(formimino)等基团的载体,在体内很多重要物质的合成中起作用。例如 N^5,N^{10}- 甲川基四氢叶酸和 N^{10}- 甲酰四氢叶酸参与嘌呤核苷酸的合成,其中甲川基和甲酰基分别成为嘌呤环中第 8 位和第 2 位上两个碳原子的来源。在脱氧尿苷酸转变成脱氧胸苷酸的过程中,胸腺嘧啶中的甲基由 N^5,N^{10}- 亚甲基四氢叶酸提供(详见第三十二章“核苷酸代谢”)。又如 N^5,N^{10}- 亚甲基四氢叶酸提供亚甲基使甘氨酸转变成丝氨酸,N^5- 甲基四氢叶酸提供甲基使高半胱氨酸转变为甲硫氨酸(详见第三十一章“氨基酸代谢”)。

由此可见,叶酸与核苷酸以及某些氨基酸的合成有密切关系。当体内缺乏叶酸时,“一碳单位”的转移即发生障碍,核苷酸尤其是脱氧胸苷酸的合成就会减少,进而影响到骨髓中幼红细胞 DNA 的合成,使得幼红细胞的分裂速率明显下降。幼红细胞因分裂障碍而体积增大,形成巨幼红细胞(megaloblast),最终导致巨红细胞性贫血。此外已发现,许多癌细胞膜上的叶酸受体过量表达,这显然是因为癌细胞需要更多的叶酸才能满足其对核苷酸合成的大量需求。很多抗癌药物就是叶酸的类似物,其作用机制是抑制癌细胞对叶酸的利用和转化(参看第三十二章“核苷酸代谢”)。

2-氨基-4-羟基-6-甲基蝶呤　对氨基苯甲酸　寡聚谷氨酸(n=2~6)

寡聚谷氨酸叶酸

谷氨酸

5,6,7,8-四氢叶酸

图 14-6　叶酸的化学结构

N^{10}-甲酰四氢叶酸　N^5-亚胺甲基四氢叶酸　N^5-甲基四氢叶酸

N^5,N^{10}-亚甲基四氢叶酸　N^5-甲酰四氢叶酸　N^5, N^{10}-甲川基四氢叶酸

图 14-7　能被四氢叶酸转移的“一碳单位”

叶酸在绿叶中大量存在，肠道细菌也能合成它，故一般不发生缺乏病。但当吸收不良、代谢失常或细胞需要过多，以及长期使用抗生素或叶酸颉颃药的情况下，可造成叶酸缺乏。

（七）生物素

生物素又名维生素 H，由带有戊酸侧链的噻吩与尿素骈合而成（图 14-8）。

在生物体内，生物素作为多种羧化酶的辅基参与 CO_2 的固定。细胞内，在生物素蛋白质连接酶（biotin protein ligase）的催化下，生物素通过戊酸侧链与羧化酶的 Lys 残基上的 ε-NH_2 形成酰胺键。于是，生物素通过细长的碳氢链系在酶分子上，这种结构对羧化反应十分重要（参看第二十五章“糖异生”）。通常将以这种形式存在的生物素和赖氨酸残基的复合物称为生物胞素（biocytin）。

由于生物素在动、植物组织中广泛存在，肠道细菌也能合成，故一般很少发生生物素缺乏病。但是，长期生吃鸡蛋可导致该维生素的缺乏，这是因为鸡蛋清中含有一种抗生物素蛋白即亲和素，此蛋白质与生物素具有很高的亲和力，可妨碍人体对生物素的吸收。

生物素缺乏的主要症状包括鳞状皮炎、精神忧郁、脱发和无食欲等。

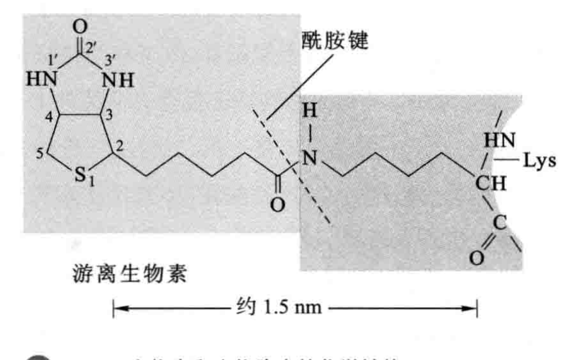


图 14-8　生物素和生物胞素的化学结构

Quiz3　鸡蛋里的抗生物素蛋白的功能是什么？在鸡蛋里还有一种蛋白质具有类似的功能，这种蛋白质是什么？

Quiz4　如何让你检测一种特定的基因是否存在，你如何设计 BAS 系统？

框 14-1　生化应用——生物素－亲和素检测系统的建立和应用

生物素分子有两个环状结构，其中咪唑酮环是与亲和素结合的主要部位；噻吩环的 C2 上有一戊酸侧链，其末端羧基可用来结合抗体和其他生物大分子，经化学修饰后，可与各种抗原、抗体、酶及核酸分子中相应基团偶联形成生物素化标记物。

亲和素是由 4 个相同亚基组成的碱性糖蛋白，每个亲和素能结合 4 个分子的生物素，二者的亲和力极强，比抗原抗体间的亲和力至少高一万倍，被认为是自然界亲和力最高的一对蛋白质－配体组合。链霉亲和素（streptavidin）则是一种由链霉菌（*Streptomyces avidinii*）分泌的蛋白质，与亲和素极为相似，也是由 4 个相同的亚基组成，每一个亚基也能结合一个生物素，但它是一种稍偏酸性的蛋白质，并且不带任何糖基。

生物素－亲和素系统（biotin-avidin-system，BAS）就是在两者高度特异性、高亲和性结合的基础上开发出的一种灵敏的生物检测系统。随着各种生物素衍生物的问世，BAS 已被广泛应用于生物学及医学研究的各个领域。生物素－亲和素与标记试剂的牢固结合及多级放大效应，使 BAS 免疫标记和有关示踪分析更加灵敏，特别适合于定性或定量检测微量抗原或抗体以及定位观察等研究。

例如，在检测一种抗原的时候，可先制备该抗原的抗体，然后将生物素与其偶联在一起，也可以将得到的抗体再次作为抗原，制备二级抗体，用生物素标记二级抗体。与此同时，将亲和素或链霉亲和素与辣根过氧化物酶偶联在一起。这样可以借助于抗原和抗体的结合，将亲和素引入到抗原所在位置，然后再借助生物素与亲和素的特异性结合，将辣根过氧化物酶引入到抗原所在位置。最后，利用辣根过氧化物酶催化的显色反应，将抗原显示出来（图 14-9）。

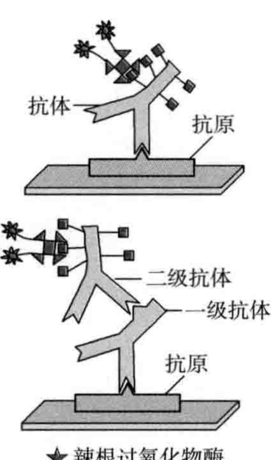


图 14-9　BAS 检测抗原图解

（八）硫辛酸

$$\text{氧化型硫辛酸} \underset{-2H}{\overset{+2H}{\rightleftharpoons}} \text{还原型硫辛酸}$$

硫辛酸　赖氨酸残基

硫辛胺复合物

图 14-10　硫辛酸的化学结构及其与硫辛酸转乙酰基酶的共价连接

硫辛酸的本质为含有 2 个硫原子的辛酸，有氧化型（2 个硫原子通过二硫键相连）和还原型（二硫键还原为巯基）两种形式，在细胞内通常以硫辛酰胺的形式与硫辛酸转乙酰酶共价相连（图 14-10），作为脂酰基的载体参与 α－酮酸的氧化脱羧（详见第二十三章“三羧酸循环”）。连接方式与生物胞素相似，也是通过羧基与酶的赖氨酸残基上的 ε-NH_2 形成酰胺键。

游离的还原型硫辛酸还是一种超强的抗氧化剂（antioxidant），可以帮助机体清除自由基，以及保护和再生其他抗氧化剂，如维生素 C 和 E。

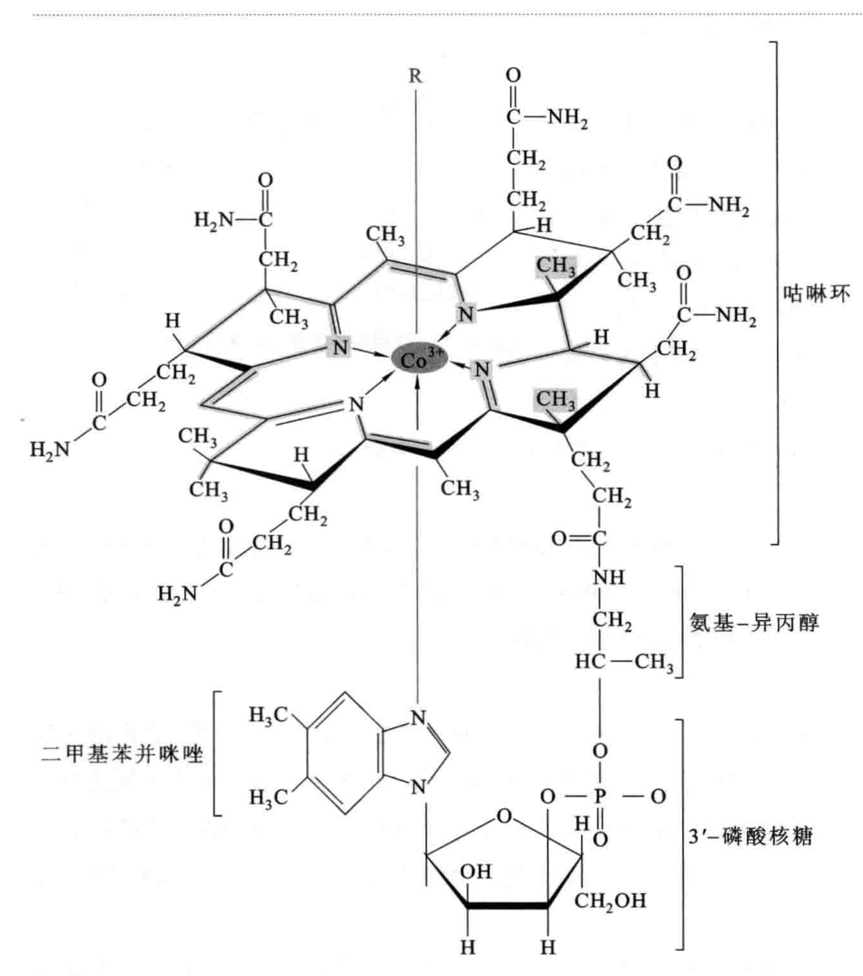
硫辛酸在自然界分布广泛，肝和酵母中的含量尤为丰富，人类还没有发现与硫辛酸相关的缺乏病。

（九）维生素 B_{12}

维生素 B_{12} 含有复杂的咕啉环结构（类似于卟啉环），可谓是自然界最复杂的辅助因子，因其分子中含有金属元素钴和若干酰胺基，故又称为钴胺素。

维生素 B_{12} 特有的结构特征是含有活泼的 C—Co 键，一旦这个化学键断裂，将产生自由基，从而引发催化。分子中的钴可以是一价、二价或三价的，并能与—CN、—OH、—CH_3 或 5′－脱氧腺苷等基团相连，分别称为氰钴胺素、羟钴胺素、甲基钴胺素和 5′－脱氧腺苷钴胺素（图 14-11）。其中甲基钴胺素和 5′－脱氧腺苷钴胺素为维生素 B_{12} 的两种辅酶形式，但两者在代谢中的作用并不同。

甲基钴胺素（CH_3—B_{12}）参与体内的转甲基反应和叶酸代谢，是 N^5－甲基四氢叶酸甲基转移酶的辅酶。此酶催化 N^5－甲基四氢叶酸和高半胱氨酸之间发生不可逆的甲基移换反应，产生四氢叶酸和甲硫氨酸（参看第三十一章“氨基酸代谢”）。



R	名称
—CN	氰钴胺素
5′-脱氧腺苷	脱氧腺苷钴胺素
—OH	羟钴胺素
$—CH_3$	甲基钴胺素

图 14-11 维生素 B_{12} 及其衍生物的化学结构

$$\underset{\text{谷氨酸}}{\begin{array}{l}COO^-\\ |\\ CH-\overset{+}{N}H_3\\ |\\ CH_2\\ |\\ CH_2\\ |\\ COO^-\end{array}} \xrightarrow[\text{5′-脱氧腺苷钴胺素}]{\text{甲基天冬氨酸变位酶}} \underset{\text{甲基天冬氨酸}}{\begin{array}{l}COO^-\\ |\\ CH-\overset{+}{N}H_3\\ |\\ CH-COO^-\\ |\\ CH_3\end{array}}$$

图 14-12 5′-脱氧腺苷钴胺素作为辅酶参与的反应

5′-脱氧腺苷钴胺素(5′-dA-B_{12})在体内作为几种变位酶的辅酶,参与反应的是活泼的 C—Co 键,例如甲基天冬氨酸变位酶(图 14-12)。

维生素 B_{12} 与叶酸一样参与体内“一碳单位”的代谢,因此它与叶酸的作用常常相互关联。缺乏维生素 B_{12} 的临床表现有恶性贫血(pernicious anemia)和神经系统受损。

维生素 B_{12} 只能由某些细菌和古菌合成,少见于植物,但可贮存在动物性食品中,特别是肝。人体对它的需要量甚少(每日仅需 2 μg),肠道细菌也能合成它,因此因摄入不足而导致维生素 B_{12} 缺乏在临床上比较少见,但严格的素食者还是有可能会缺乏维生素 B_{12} 的。此外,维生素 B_{12} 的吸收与胃黏膜分泌的一种糖蛋白密切相关,这种糖蛋白称为内在因子(intrinsic factor,IF)。维生素 B_{12} 必须与内在因子结合后才能被小肠吸收。若病人患有萎缩性胃炎、胃全切除或先天缺乏 IF,均可出现维生素 B_{12} 的缺乏。

二、维生素 C

维生素 C 又名抗坏血酸(ascorbic acid),是含有内酯结构的酸性多羟基化合物,其分子中第 2 位和第 3 位碳原子上的两个烯醇式羟基极易解离质子,因而在水溶液中有较强的酸性。此外,维生素 C 可氧化成脱氢维生素 C,此反应是可逆的(图 14-13)。维生素 C 含有手性碳原子,因而具有光学异构体,自然界存在的具有生理活性的是 L 型。

许多动物能够利用葡萄糖作为前体合成维生素 C,但是灵长类、某些鸟类、鱼类、无脊椎动物和豚

鼠不行,原因是它们体内缺少合成维生素 C 的一个关键酶。

维生素 C 除了防治坏血病(scurvy)外,在临床上还有其他用途。从防感冒、消炎、防止动脉硬化到抗癌,似乎有百利而无一弊。Pauling 曾公开宣称,高剂量的维生素 C 可以预防感冒和减轻感冒的症状,并写了一本名为《维生素 C 与普通感冒》的书。可以说维生素 C 是目前应用最广的一种维生素,使用的剂量越来越大,似乎已被作为保健药物来使用了。

L-抗坏血酸　　L-脱氢抗坏血酸

图 14-13　还原型维生素 C 和氧化型维生素 C 的互变

已知维生素 C 在体内的生理功能是多方面的,主要有以下两个方面。

1. 参与体内的羟基化反应

维生素 C 在细胞内参与多种物质的羟基化反应,而羟基化反应又是体内许多重要化合物的合成或分解的必经步骤。例如,胶原蛋白的后加工、类固醇的合成与转变、胆酸的形成、肾上腺素的合成、芳香族氨基酸的羟基化以及许多有机药物或毒物的生物转化。

Quiz5　临床上使用大剂量维生素 C 可降低血中胆固醇,为什么?

2. 抗氧化作用

(1) 保护肽和酶分子上游离的巯基、促进巯基的再生　已知许多含巯基的酶要依赖于自由的巯基(—SH)才能有活性,而维生素 C 可以防止酶分子中的—SH 被氧化。此外,维生素 C 有助于氧化型的谷胱甘肽(G—S—S—G)还原成还原型的谷胱甘肽(G—SH),使—SH 得以再生,从而保证谷胱甘肽的功能。某些含巯基的酶在重金属中毒(如铅中毒)时被抑制,补充大量维生素 C 往往可以缓解毒性,其原理就在此。

(2) 防止铁的氧化、促进铁的吸收　维生素 C 能使难吸收的 Fe^{3+} 还原成易吸收的 Fe^{2+},从而促进铁的吸收。它还能促使体内的 Fe^{3+} 还原,促进血红素的合成。此外,维生素 C 还有直接还原高铁血红蛋白的作用。

Quiz6　哪些水溶性维生素衍生出的辅酶或辅基含有腺苷或腺苷酸?这样的性质对研究生命的进化有什么意义?

第二节　脂溶性维生素

脂溶性维生素包括维生素 A、维生素 D、维生素 E 和维生素 K(为了便于记忆,可将它们拼写成 DAKE),均是异戊二烯衍生物。

一、维生素 A

维生素 A 是由 β- 白芷酮环和两个异戊二烯单位缩合而成的不饱和一元醇,有 A_1 和 A_2 两种(图 14-14)。A_1 为视黄醇(retinol),A_2 为 3- 脱氢视黄醇。A_1 在体内经脱氢可转变为 11- 顺视黄醛(11-*cis*-retinal),11- 顺视黄醛可异构化为全反式视黄醛(all-*trans*-retinal)。11- 顺视黄醛可进一步被氧化成视黄酸(retinoic acid),但此反应是不可逆的。

Quiz7　何种酶催化视黄醇转变为 11- 顺视黄醛?

维生素 A 的生理功能由视黄醇、视黄醛和视黄酸来完成,主要表现在以下四个方面。

(1) 视黄醇和视黄酸可作为脂溶性激素,通过与它们的细胞核受体的结合来启动某些基因的表达(参看第十七章“激素及其受体介导的信号转导”),从而促进细胞的生长和分化,还可以阻止角蛋白的合成,并具有一定的抗癌作用。与此同时,视黄醇的磷酸酯——磷酸视黄醇(retinyl phosphate)作为糖基供体直接参与某些糖蛋白和粘多糖的合成,这些糖蛋白和黏多糖绝大多数是上皮组织分泌黏液的主要成分,参与调节细胞的生长。由此看来,维生素 A 是维持一切上皮组织的完整与健全所必需的物质,缺乏时上皮干燥、增生及角蛋白大量分泌导致角质化,其中对眼部、消化道、呼吸道、

图 14-14　β- 胡萝卜素向维生素 A 的转变以及维生素 A 在体内的功能

尿道、膀胱及生殖系统等处的上皮影响最为显著。在眼部，由于泪腺上皮角化，泪液分泌受阻，以致角膜、结膜干燥而产生干眼病（xerophthalmia），因此维生素 A 又称为抗干眼病维生素。

（2）视黄醛构成视网膜的感光物质，作为视蛋白（opsin）的辅基参与视觉的形成（详见第十七章“激素与受体介导的信号转导”），缺乏它可导致夜盲症（night blindness）。

（3）维生素 A 的抗氧化作用（参看维生素 C 和维生素 E）比较明显。维生素 A 只存在于动物性食品中，但是在很多植物性食品如胡萝卜、红辣椒、菠菜、芥菜等有色蔬菜中含有维生素 A 的前体——β-胡萝卜素（β-carotene）。β- 胡萝卜素可被小肠黏膜或肝中的一种酶裂解为视黄醇，故又被称作维生素 A 原（provitamin A）。

二、维生素 D

维生素 D 属于固醇类衍生物（图 14-15）。人体内的维生素 D 主要是由 7- 脱氢胆固醇经紫外线照射转变而成，也可从动物食品中获取，这种形式的维生素 D 被称为维生素 D_3 或胆钙化醇（cholecalciferol）。植物中的麦角固醇经紫外线照射后可产生维生素 D_2 或钙化醇。

7- 脱氢胆固醇存在于皮肤内，它可由胆固醇脱氢产生。一般人体只要充分接受阳光照射，通过这种方式合成的维生素 D 完全可以满足生理需要。

两种维生素 D 的生理作用基本相同，但不论是哪一种，本身都没有明显的生理活性，它们必须先在肝细胞内经羟基化转变为 25- 羟基维生素 D，然后在肾小管内进行第二次羟基化反应，最后形成具有活性的 1,25- 二羟维生素 D［1,25-$(OH)_2D$］。1,25-$(OH)_2D$ 作为一种脂溶性激素发挥作用，可诱

图 14-15　维生素 D 的化学结构

Quiz8 肾功能不全的人如何补充维生素 D?

导某些基因表达，如骨钙蛋白(osteocalcin)。

维生素 D 在体内与甲状旁腺素协同作用，共同促进小肠对食物中钙和磷的吸收，维持血中钙和磷的平衡，促进骨和齿的钙化作用。由于维生素 D 具有抗佝偻病(rickets)的作用，故又名抗佝偻病维生素。

三、维生素 E

维生素 E 又称为生育酚，有 α、β、γ 和 δ 四种，其中以 α- 生育酚(图 14-16)的生理效用最强。它们都可视为苯骈二氢吡喃的衍生物。

图 14-16 α- 生育酚的化学结构

维生素 E 的主要生理功能是在体内作为一种强抗氧化剂，与硫辛酸、维生素 A、β- 胡萝卜素、维生素 C、超氧化物歧化酶、过氧化氢酶及谷胱甘肽过氧化物酶一起，防止自由基和过氧化物对脂质(特别是不饱和脂肪酸)的氧化，保护细胞膜免受氧化损伤以及维护红细胞的完整。其次，维生素 E 在体内还可以去调节某些酶的活性，例如在抑制蛋白质激酶 C(PKC)的活性的同时，激活磷蛋白磷酸酶 2A 的活性，从而影响到平滑肌的生长。

此外，维生素 E 还参与生物氧化，在呼吸链中既可以稳定辅酶 Q，又可以协助电子传递给辅酶 Q。维生素 E 还与动物生殖机能有关，因此又名为生育酚。

维生素 E 分布极广，在植物油中特别丰富，一般不易缺乏。

四、维生素 K

维生素 K 是 2- 甲基 1,4- 萘醌的衍生物(图 14-17)，自然界已发现的有存在于绿叶植物中的维生素 K_1 和由肠道细菌合成的维生素 K_2。

维生素K_1

维生素K_2

图 14-17 维生素 K 的化学结构

在体内，维生素 K 主要作为依赖于维生素 K 的羧化酶(vitamin K-dependent carboxylase)的辅酶，去参与某些蛋白质的后加工，使这些蛋白质分子上特定的谷氨酸残基经历 γ- 羧基化修饰，并最终激活它们的活性(图 14-18)。需要进行 γ- 羧基化修饰的蛋白质有凝血因子(clotting factor)Ⅱ、Ⅶ、Ⅸ和Ⅹ以及骨钙蛋白。其中凝血因子可以促进血液凝固，因此维生素 K 又名为凝血维生素。而骨钙蛋白能够结合钙离子，因此维生素 K 也参与骨的形成。此外，在某些生物体内，维生素 K 还可以作为呼吸链的一部分，参与生物氧化。

Quiz9 哪些维生素人体也可以合成一点，哪些维生素肠道细菌可以为我们合成一点?

CO_2 → 依赖于维生素K的羧化酶；谷氨酸残基 → γ-羧基谷氨酸残基 → Ca^{2+}

图 14-18 维生素 K 作为辅酶参与的酶促反应

因为肠道细菌能合成维生素 K，所以人类维生素 K 缺乏病多为吸收障碍或因长期使用抗生素或维生素 K 的代谢颉颃药（metabolic antagonists）所致。

科学故事——维生素的发现

在第一种维生素（VB_1）被发现之前，许多特定食物的一些特殊的防病、治病的作用就早已被人们知晓。中国唐代医学家孙思邈曾经指出，使用动物肝可以防治夜盲症，用谷皮熬粥可以防治脚气病。这些食物中真正起作用的成分正是维生素，例如，动物肝中含有丰富的维生素 A，而谷皮中含有丰富的维生素 B_1，它们分别是治疗夜盲症和脚气病的特效药。

1886 年，年轻的荷兰军医 Christiaan Eijkman 前往 Java 岛上的一个研究所工作，研究亚洲普遍流行的脚气病。前几年在研究所的工作中，他的两个同事想方设法从死于脚气病的病人体内提取致病的微生物，却一直没有成功。当他们回到欧洲的时候，工作就留给了 Eijkman。

Eijkman 尝试用微生物感染兔子和猴子，但是，动物并没有得病。于是，Eijkman 得出结论：脚气病一定需要长时间才能发病。为此，他等了很久。可是无论是兔子，还是猴子就是不得病。相反，这些动物患其他疾病的机会增大了。看来要想让动物得脚气病，必须同时饲养大量的动物，但还要考虑成本和是否容易饲养等因素。

于是，他购买了鸡，将它们关在研究所屋顶下面搭建的一个大的鸡笼里饲养。在不到一个月的时间内，所有的鸡都得病了！ Eijkman 认为是被注射了细菌的鸡患病后传给了没有注射的鸡。为了验证他的结论，他又购买了一批鸡。这次是将每一只鸡分开来喂养。 然而，这些鸡也患病了。Eijkman 意识到，一定是整个研究所都被细菌感染了，于是决定将新买的鸡放到新的地方去饲养。在这样做以后，他发现所有的鸡都好了。Eijkman 不知道究竟是什么使鸡痊愈的！ 这时，那个喂鸡的饲养员告诉 Eijkman，在鸡患病期间，他喂给鸡的是来自于隔壁医院剩下来的精白米。后来，那里一个新来的厨师不给他们剩的白米了，他才重新给鸡喂糙米。就在这以后，鸡痊愈了。

根据饲养员的叙述，Eijkman 推测鸡的脚气病与饲料有关。为此，他决定做下面一些实验以进一步确认他的推测：他同时饲养了 4 组共 11 只鸡，有 3 组喂糙米，一组喂精白米。在 5 个星期以后，Eijkman 发现，只有喂精白米的鸡才患病，而吃糙米的鸡都很健康，而且原来患病的鸡也痊愈了！ 随后，他给所有的鸡喂糙米，结果 4 只病鸡也都痊愈了。这就证实了鸡饲料与脚气病之间有关联的推测。

不久，Eijkman 重复了实验，以确认病因的确在于烹调用的精白米上。对此，他提出了各种可能的解释：是不是精白米帮助了一种未知的微生物在鸡肠道内致毒的？ 或者只是白米放久了而有毒？ 为了测试后一种可能性，他给鸡吃新鲜的白米，但鸡仍然生病了。那么是不是鸡不能从精白米中吸收营养呢？ 或者煮白米的水有毒呢？

结果所有的答案都是否，但有一件事很明确，就是喂给鸡吃的米有没有米糠（谷皮）的关系重大。只喂养精米的鸡才得病。如果鸡喂给精米，再添加米糠，它们就不得病。后来，Eijkman 给鸡喂生肉，结果鸡都很健康。Eijkman 因此得出结论：淀粉里面含有某些毒性成分，而在白米中被去掉的米糠中含有一种使毒性成分无害的物质，并称之为抗脚气病因子。

1895 年，在对动物进行 9 年的研究以后，Eijkman 想调查人是否能通过吃糙米的方式来避免患脚气病。他请了一个名叫 A. G. Vorderman 的医生在监狱里进行研究。选择监狱的原因是因为在那种地方很容易控制饮食的构成，而且相同的人可以待在那里很长的时间。结果很快发现，吃精米的犯人更容易得脚气病。在这个实验完成以前，Eijkman 回到了荷兰。Vorderman 和其他几名医生继续他们的研究。在 Eijkman 离开研究所以后，Gerrit Grinjs 接管了他的职位，并做了更多的研究。根据他的研究，Grinjs 认为，食物中含有人需要的其他东西，白米也不是有毒的，它只是缺乏某种重要的成分。

1906 年，英国生物化学家 Frederick Growland Hopkins 用纯化后的精饲料喂食老鼠，饲料中含有蛋白质、碳水化合物、脂类和矿物质，然而老鼠依然不能存活；而如果向纯化后的饲料中加入哪怕只是微量的牛奶后，老鼠就能正常生长了。从而证明食物中除了蛋白、糖类、脂类、微量元素和水等营养物质外还存在一种必需的“辅助因子”（accessory factor）。

1911年，波兰化学家 Casimir Funk 发现糙米中能够防治脚气病的物质是一种胺，因此提议将这种化合物叫做 Vitamine，意为"Vital amine"，其意思就是"维持生命所必需的胺"。这个名词迅速被用于其他"辅助因子"。但很快就发现，许多其他的"辅助因子"并不含有"胺"结构，但是由于 Funk 的叫法已被广泛采用，因而 Vitamine 这种叫法并没有被废弃，而仅仅将 amine 的最后一个"e"去掉，改成了"vitamin"。

1912年，Hopkins 和 Funk 推出维生素缺乏假说，推测人体中如果缺乏特定的足够量的维生素，将会引起特定的疾病。Funk 并没有发现真正的维生素物质。直到1926年，后来被命名为硫胺素或 B_1 的维生素才被纯化。到了1936年，它的结构才被阐明，并最终被人工合成。1929年，Eijkman 与 Hopkins 共同荣获了诺贝尔生理学或医学奖。

e14-2 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Vitamin（维基百科有关维生素的内容）
2. http://themedicalbiochemistrypage.org/vitamins.html（完全免费的医学生物化学课程网站有关维生素的内容）
3. http://lpi.oregonstate.edu/infocenter/vitamins.html（美国俄勒冈州立大学附属的 Linus Pauling 研究所有关维生素的网站，内有许多有关维生素的文献）

参考文献：

1. Sanjoaquin M A, et al. Folate intake and colorectal cancer risk: a meta-analytical approach. *Int J Cancer*, 113(5): 825–828, 2005.
2. Jacobson M K, Jacobson E L. Discovering new ADP-ribose polymer cycles: protecting the genome and more. *Trends Biochem Sci*, 24(11): 415–417, 1999.
3. Furie B, et al. Vitamin K-dependent biosynthesis of gamma-carboxyglutamic acid. *Blood*, 93(6): 1 798–1 808, 1999.

第十五章　糖类

糖类是指多羟基醛或多羟基酮以及它们的缩合物和某些衍生物。含有醛基的糖称为醛糖(aldoses)，含有羰基的糖称为酮糖(ketoses)。由于最早发现的几种糖类化合物可以用通式 $C_n(H_2O)_m$ 来表示，因此糖类又称为碳水化合物(carbohydrates)。根据聚合度的不同，糖类可以分为单糖(monosaccharides)、寡糖(oligosaccharides)和多糖(polysaccharides)。

糖类是自然界含量最丰富的有机分子，在生物体内的主要功能包括：氧化放能产生 ATP；充当生物贮存能量的一种方式；作为合成其他生物分子的前体和细胞的结构组分；参与细胞与细胞之间的分子识别和信号转导等。

本章将主要介绍糖类的结构、性质和功能，重点是单糖的结构和性质，特别是它们的立体结构。

第一节　单糖

一、单糖的命名和缩写

单糖又名简单糖(simple sugars)，已不能再水解成更简单的糖单位。根据碳原子的数目，可将单糖分为丙糖(triose)、丁糖(tetrose)、戊糖(pentose)、己糖(hexose)和庚糖(heptose)等，它们含有的碳原子数目分别是 3、4、5、6 和 7 等。

根据各单糖的化学结构，丙糖以外的单糖可看成是由丙糖衍生而来的，其中醛糖衍生于甘油醛，酮糖衍生于二羟丙酮。

具体的单糖多是根据各自的来源来命名的，如葡萄糖(glucose)、果糖(fructose)和半乳糖(galactose)。

可以使用三字母缩写(通常是英文单词的前 3 个字母)表示一种单糖，而相应的衍生物的缩写则可能不止三个字母，注意首字母要大写。

e15-1 常见单糖的名称及其英文缩写

二、单糖的旋光异构

最简单的单糖是丙糖，包括甘油醛(glyceraldehyde)和二羟丙酮(dihydroxyacetone)。其中二羟丙酮无手性 C，而甘油醛含有 1 个手性 C，所以甘油醛分为 D 型和 L 型。这两种甘油醛呈镜像关系，属于对映异构体(enantiomers)或镜像异构体。根据惯例，在甘油醛的 Fischer 投影结构式中，醛基画在最上方，羟基位于左侧的甘油醛定为 L 型，羟基位于右侧的甘油醛定为 D 型(图 15-1)。

图 15-1　甘油醛和二羟丙酮的 Fischer 投影结构式

除了二羟丙酮以外，其他单糖至少含有一个手性 C，因此都具有旋光异构体。为了方便，通常以 D-甘油醛和 L-甘油醛为标准，将单糖分为 D-单糖和 L-单糖。与氨基酸正好相反，自然界的单糖绝大多数为 D-型。

要是一种单糖含有的手性 C 不止一个，那么在确定它究竟是 D 型还是 L 型的时候，需要将它的 Fischer 投影结构式之中编号最高的手性 C(距离羰基最远的手性 C)与甘油醛的手性 C 进行比较，与 D 型甘油醛一致的就是 D 型，反之就是 L 型。D 型和 L 型并不能提供任何有关旋光方向的信息。

在各种旋光异构体之中，除了互为镜像的对映异构体以外(图 15-2)，还有非对映异构体和差向异构体两种情形。如果一对旋光异构体有一个或一个以上的手性 C 的构型相反，但并不呈镜像关系，那么就称为非对映异构体(diastereomers)；如果一对旋光异构体只有一个手性 C 的构型不同，则称为差向异构体。例如 D-葡萄糖与 D-甘露糖，D-葡萄糖与 D-半乳糖就互为差向异构体(epimers)

图 15-2　果糖的对应异构体

图 15-3　D- 葡萄糖的差向异构体

(图 15-3),但 D- 甘露糖和 D- 半乳糖并非差向异构体,只能是非对映异构体,因为它们之间有两个手性碳的构型不同。

三、单糖的环状结构和异头物

如图 15-4 所示,醇羟基很容易与醛或酮形成半缩醛(hemiacetal)或半缩酮(hemiketal)。除了丙糖和丁酮糖以外,其他直链的单糖分子在分子内也能发生类似的反应,形成环状结构,其中醛糖环化形成环式半缩醛,酮糖环化形成环式半缩酮。丙糖和丁酮糖之所以不能形成环状的结构,是因为如果它们发生分子内的半缩醛或半缩酮反应的话,只能形成三元环或四元环,而这样的环结构是很难稳定存在的。

实验证明,葡萄糖环化主要形成六元环的吡喃糖(pyranose),果糖、核糖和脱氧核糖环化主要形成五元环的呋喃糖(furanose)(图 15-5 和图 15-6)。

通常使用 Haworth 式来表示单糖的环状结构。按照如图 15-7 所示的方法,可以很容易地将一种

图 15-4　缩醛和缩酮反应

图 15-5　吡喃葡萄糖和呋喃果糖

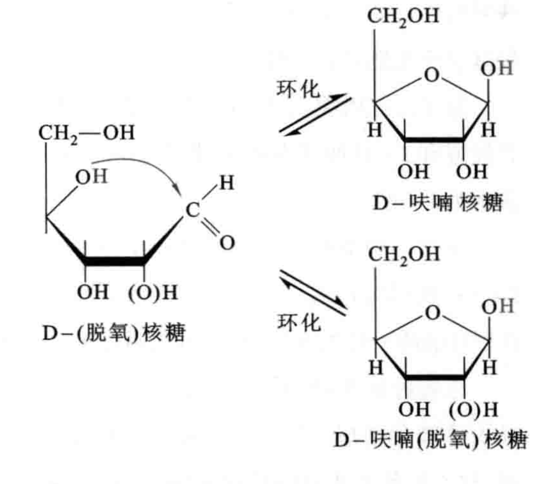


图 15-6　(脱氧)核糖的环化

单糖的开链 Fischer 投影式正确地过渡到相应的环状 Haworth 式。

然而，一旦单糖由直链结构变成环状结构以后，原来的羰基 C 便“摇身一变”，成为一个新的手性中心，从而产生 α 和 β 两种异构体。这种在半缩醛 C 上形成的异构体称为异头物（anomers），新出现的手性 C 称为异头物 C（anomeric carbon）。如图 15–8 所示，如果半缩醛羟基与原来编号最高的手性 C 上的羟基具有相同的取向，这种异头物就称为 α 异头物，反之就称为 β 异头物。由于 β 异头物比 α 异头物稳定，因此在葡萄糖溶液之中，β–D– 葡萄糖要比 α–D– 葡萄糖多。

Quiz1 你认为 α–D– 葡萄糖和 β–D– 葡萄糖是一对差向异构体吗？

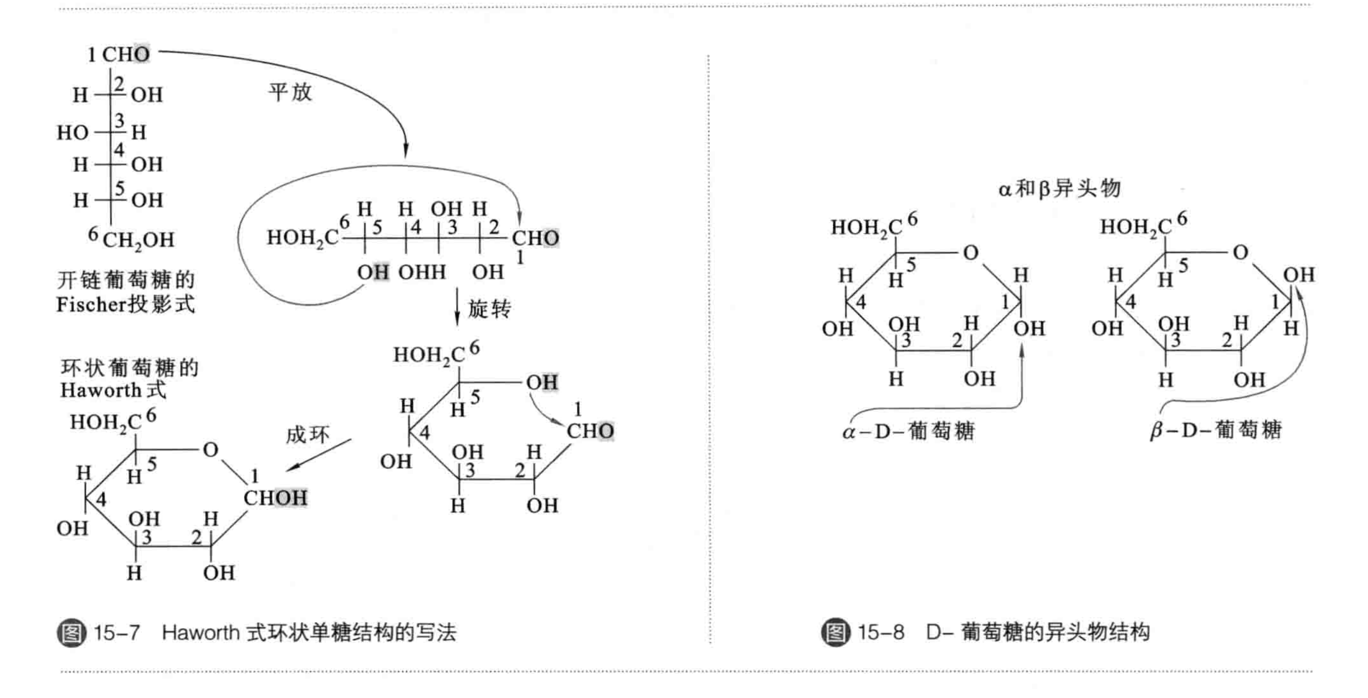


图 15–7 Haworth 式环状单糖结构的写法

图 15–8 D– 葡萄糖的异头物结构

四、单糖的构象

以葡萄糖为例，其半缩醛环上的 C—O—C 键角为 111°，与环己烷的键角（109°）相近，故葡萄糖的吡喃环和环己烷环相似，也有椅式构象（chair conformation）和船式构象（boat conformation），其中椅式构象可使各单键的扭张强度降低到最小因而较稳定。而在图 15–9 所示的两种椅式结构之中，Ⅰ型上的—OH 和—CH_2OH 这两种较大的基团均为平伏键，将可能产生的空间位阻降低到最小，所以在热力学上Ⅰ型比Ⅱ型稳定。

(1) 轴 椅式 轴 船式

a=直立键
e=平伏键

(2) 椅式Ⅰ(稳定) 环转向 椅式Ⅱ(不稳定)

图 15–9 葡萄糖的椅式构象和船式构象

五、单糖的衍生物

在特定的酶催化下，单糖在体内可进行各种修饰反应而形成一系列衍生物。常见的衍生物包括氨基糖（amino sugar），如葡糖胺（glucosamine）、*N*-乙酰葡糖胺（*N*-acetyl glucosamine，NAG）、半乳糖胺（galactosamine）、*N*-乙酰氨基半乳糖胺、胞壁酸（muramic acid）和神经氨酸（neuraminic acid）；氧化糖，如葡糖内酯、葡糖酸（gluconic acid）、葡糖醛酸（glucuronic acid）、半乳糖醛酸（galacturonic acid）和甘露糖醛酸（mannuronic acid）；脱氧糖（deoxy-sugar），如2-脱氧核糖、L-鼠李糖（L-rhamnose）和L-岩藻糖（L-fucose）；糖醇（alditol），如山梨醇（sorbitol）、甘露醇（mannitol）、木糖醇（xylitol）、半乳糖醇（galactitol）、核糖醇（ribitol）和肌醇（innositol）；糖苷（glucoside），如毛地黄苷（digitoxin）和乌本苷（oubain）。

e15-2 各种单糖衍生物的化学结构

六、单糖的性质

（一）物理性质

e15-3 单糖能发生的主要反应

室温下的纯单糖多为无色晶体，一般具甜味，溶于水，微溶于乙醇。除了二羟丙酮，其他所有的单糖都具有旋光性。

（二）化学性质

单糖具有羰基和多个羟基，能与多种化学试剂反应。

1. 变旋

变旋（mutarotation）是指一种碳水化合物的两种异头物在水溶液中发生互变，并达到平衡，从而导致旋光度改变的现象。例如，新鲜配制的α-D-葡萄糖溶液或β-D-葡萄糖的比旋光度分别是112.0和18.7，变旋后的比旋光度均为52.7。

Quiz2 有一种刚配制好的单糖溶液，开始是甜的，但过一段时间后变苦了，对此试给予解释。

变旋的直接原因是环状的异头物转变成开链形式，而当开链形式重新变为异头物的时候有两种可能，即可能变为α异头物或β异头物。在任何情况下，只有很少一部分为开链形式。如果一种单糖不能形成环状结构，就不会有变旋现象，如甘油醛。

2. 异构

单糖分子中，α-氢原子由于受到羰基和羟基的双重影响而变得十分活泼，在碱性条件下，醛糖可通过烯二醇（enediol）式中间物与2-酮糖实现互变（图15-10）。

Quiz3 D-葡萄糖溶液里究竟有多少种异构体？

D-葡萄糖、D-甘露糖和D-果糖在C3、C4、C5上的构型相同，因此它们具有相同的烯二醇中间物，在碱性溶液中，实际上存在着这三种糖的平衡。

图 15-10 酮糖和醛糖的互变

3. 成脎反应

Quiz4 能否利用成脎反应来区分D-葡萄糖、D-果糖和D-甘露糖？

所有还原糖的C1和C2都能够与苯肼起反应生成糖脎（osazone）。糖脎的黄色结晶十分有用，可用来鉴别还原糖的存在，还可以从糖脎的熔点和晶形来鉴别某些糖。

4. 成苷反应

糖苷是糖分子中的半缩醛或半缩酮羟基与另一分子羟基或氨基化合物脱水生成的缩醛或缩酮结构的化合物。糖苷分子中糖的部分称为糖基，非糖部分称为配基，连接糖基和配基的键叫做糖苷键。例如，β–D–葡萄糖与甲醇在盐酸的催化下，形成甲基 –β–D–葡糖苷。

由 α– 型半缩醛或半缩酮羟基所形成的糖苷称作 α– 糖苷，其糖苷键称作 α– 糖苷键；由 β– 型半缩醛或半缩酮羟基所形成的糖苷称作 β– 糖苷，其糖苷键称作 β– 糖苷键。自然界中以 β– 糖苷为主，存在于植物的根、茎、花叶和种子中。核酸分子中的核苷实际上是由碱基与核糖或脱氧核糖之间形成的 β– 糖苷。

与单糖不同的是，糖苷虽然仍有旋光作用，但因半缩醛结构不复存在而无法开环，所以无变旋现象；糖苷不能被 Fehling 试剂或 Benedict 试剂所氧化，且不能成脎；糖苷在酸性条件下能发生水解反应，但在碱性条件下能稳定存在。

5. 呈色反应

(1) Molish 反应　糖类化合物与 α– 萘酚 / 乙醇在试管中混合，摇匀后沿管壁滴加浓硫酸，在两液面交界处出现紫红色环，此反应称为 Molish 反应。使用此反应，可以将糖类与非糖类化合物区分开。

(2) Seliwanoff 反应　糖类化合物与浓酸作用后再与间苯二酚反应，若是酮糖就显鲜红色，若是醛糖就显淡红色，这种反应称为 Seliwanoff 反应。根据此反应可鉴别酮糖和醛糖。

(3) 间苯三酚反应　戊糖与间苯三酚 / 浓盐酸反应生成朱红色物质，其他单糖与间苯三酚 / 浓盐酸生成黄色物质；此外，戊糖还可以和甲基间苯二酚即地衣酚 / 浓盐酸反应，生成蓝绿色物质。利用这两个反应可以将戊糖和其他单糖区分开来。

除了以上 5 种反应以外，单糖还可以发生氧化反应或还原反应。其中所有的单糖都能与弱氧化剂起反应，如 Fehling 试剂或 Benedict 试剂，因此所有的单糖都具有还原性。

e15–4 单糖的氧化和还原反应

七、几种生化上重要的单糖

(1) D– 甘油醛和二羟丙酮　两者都是糖酵解的中间物（参看第二十二章“糖酵解”）。

(2) D– 葡萄糖　是细胞重要的能源，对于神经细胞尤为重要。血糖由溶解在血液中的葡萄糖组成，其浓度需要稳定在 0.7~1 mg/ml。

(3) D– 半乳糖　在动物体内较为少见，它通常是脑和神经组织中糖蛋白的成分，因此有时被称为脑糖（brain sugar）。

(4) D– 果糖　主要存在于许多果实和蜂蜜之中，在所有碳水化合物中甜度最高。

(5) D– 核糖和 2– 脱氧 –D– 核糖　两者分别是核糖核苷酸（RNA）和脱氧核苷酸（DNA）的组分。

第二节　寡糖

寡糖也称为低聚糖，由 2~10 个单糖分子缩合并以糖苷键相连。在寡糖分子之中，异头物 C 上的半缩醛羟基以游离形式存在的一端为还原端（reducing end），异头物 C 参与形成糖苷键的一端为非还原端（non-reducing end）。

按照惯例，在书写寡糖序列时，非还原端写在左边，还原端写在右边，同时需要标明各单糖单位的名称、构型、相互间的连接方式和异头物的构型。

二糖（disaccharide）是最简单的寡糖，其中的一个单糖单位的连接碳总是 C1，而另一个单糖单位的连接 C 的位置是可变的。在某种意义上，二糖可视为由两分子单糖连接而成的糖苷。与其他糖苷一样，二糖在酸性条件下糖苷键发生水解，生成两分子单糖。

Quiz5 假定每一个单糖都以吡喃环形式存在，那么将一个葡萄糖、一个甘露糖和一个半乳糖连接起来可以形成多少种不同的寡糖？比较由 3 种不同的氨基酸形成多少种不同的 3 肽。

根据能否被弱氧化剂(如 Fehling 试剂)氧化的性质,二糖可以分为还原性和非还原性两类。常见二糖的名称、结构、来源和生理功能参见表 15–3。

► 表 15–3　常见二糖的名称、结构、来源和生理功能

二糖	性质	结构	来源	生理功能
蔗糖	非还原糖	Glcα(1→2)Fruβ	水果、种子、根和蜂蜜	光合作用的终产物
乳糖	还原糖	Galβ(1→4)Glc	动物乳汁、某些植物	动物的能源
α,α– 海藻糖	非还原糖	Glcα(1→1)Glcα	藻类、真菌和昆虫血	昆虫血糖
麦芽糖	还原糖	Glcα(1→4)Glc	淀粉和糖原	淀粉和糖原中的二糖单位
异麦芽糖	还原糖	Glcα(1→6)Glc	支链淀粉和糖原	淀粉和糖原分支
纤维二糖	还原糖	Glcβ(1→4)Glc	植物	纤维素中的二糖单位
龙胆二糖	还原糖	Glcβ(1→6)Glc	某些植物(如龙胆属)	植物的糖苷组分

1. 还原性二糖

还原性二糖是由一分子单糖的半缩醛羟基与另一分子单糖的醇羟基缩合而成,例如麦芽糖(maltose)、乳糖(lactose)。由于还原性二糖分子中还存在着一个游离的半缩醛羟基,因此它们不但具有变旋现象,而且能够被 Fehling 试剂或 Benedict 试剂氧化。

2. 非还原性二糖

非还原性二糖是由二分子单糖的半缩醛羟基脱水而成的,最常见的是蔗糖和海藻糖。非还原二糖具有旋光性,但由于不存在游离的半缩醛羟基,因此既无变旋现象,也不能被 Fehling 试剂或 Benedict 试剂氧化,更不能成脎。

寡糖除了能以游离的形式存在以外,还可以与蛋白质或脂质共价相连,成为糖蛋白或糖脂的重要成分,并具有多种功能(参看本章第四节糖缀合物)。

第三节　多糖

由多个单糖分子缩合而成的糖类称为多糖(polysaccharide)或聚糖(glycan),其中由相同的单糖分子组成的多糖为同多糖(homopolysaccharide),含有不同种单糖单位的多糖为杂多糖(heteropolysaccharide)(表 15–4)。多糖中最常见的单糖是 D– 葡萄糖,其次是 D– 果糖、D– 半乳糖、L– 半乳糖、D– 甘露糖、L– 阿拉伯糖和 D– 木糖。某些单糖的衍生物,例如 D– 葡糖胺、D– 半乳糖胺、*N*– 乙酰葡糖胺、*N*– 乙酰半乳糖胺、葡糖醛酸,也出现在一些多糖分子之中。

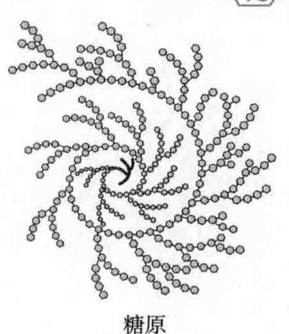
► 表 15–4　常见多糖的结构和性质

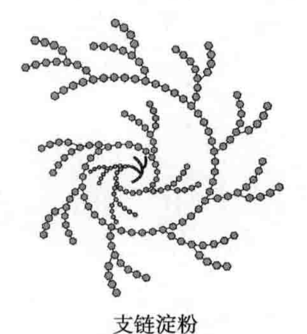
	淀粉	糖原	右旋糖酐	纤维素	几丁质	菊粉	琼脂糖
来源	种子、块茎、块根	肝脏骨骼肌	酵母细菌	植物细胞壁	无脊椎动物等	菊科植物	海藻
单糖单位	D– 葡萄糖	D– 葡萄糖	D– 葡萄糖	D– 葡萄糖	NAG	果糖、葡萄糖	D– 半乳糖、3,6– 脱水半乳糖
连接方式	α–1,4 α–1,6	α–1,4 α–1,6	主要是 α–1,6	β–1,4	β–1,4	主要是 β–2,1	β–1,4 α–1,3
类别	同多糖	同多糖	同多糖	同多糖	同多糖	杂多糖	杂多糖
碘反应	蓝色或紫红色	紫红色	无	无	无	无	无
构象	螺旋		无规卷曲	锯齿带状	带状构象		
功能	贮能	贮能	贮能	结构	结构	贮能	结构

单糖之间的连接方式，即糖苷键的类型，直接与多糖的机械强度和溶解性质有关。以 α-1，4 糖苷键相连的多糖（如淀粉和糖原）往往比较软，在水里有一定的溶解度，而以 β-1，4 糖苷键相连的多糖（如纤维素和几丁质）比较硬，难溶于水。

与蛋白质和核酸不同的是，多糖无确定的相对分子质量，因为组成多糖的单糖单位的数目不是固定的。与单糖和寡糖相比，多糖无变旋现象和还原性，也无甜味。

按照功能的不同，多糖可分为贮能多糖（storage polysaccharide）和结构多糖（structural polysaccharide）。





直链淀粉　纤维素

图 15-11　糖原、淀粉和纤维素的结构

一、贮能多糖

常见的贮能多糖有淀粉（starch）、糖原（glycogen）和右旋糖酐（dextran）。这三类多糖的基本组成单位都是 D- 葡萄糖，结构参见图 15-11，性质参见表 15-4，都属于同多糖。

生物之所以以多糖而不是单糖的形式贮存能量，是因为多糖形式能够降低糖贮备给细胞带来的渗透压（osmotic pressure）。由于渗透压只取决于分子的数目，因此，1 个游离的葡萄糖分子与 1 个由 1 万个葡糖残基组成的糖原分子给细胞带来的渗透压是相同的。

（一）淀粉

淀粉又分为直链淀粉（amylose）和支链淀粉（amylopectin）两类（图 15-12）。直链淀粉只有 α-1，4 糖苷键，无分支，其重复的二糖单位为麦芽糖，而支链淀粉既含有 α-1，4 糖苷键，又含有 α-1，6 糖苷键，其重复的二糖单位主要是麦芽糖，还有少数异麦芽糖。α-1，6 糖苷键的出现使得支链淀粉产生分支。

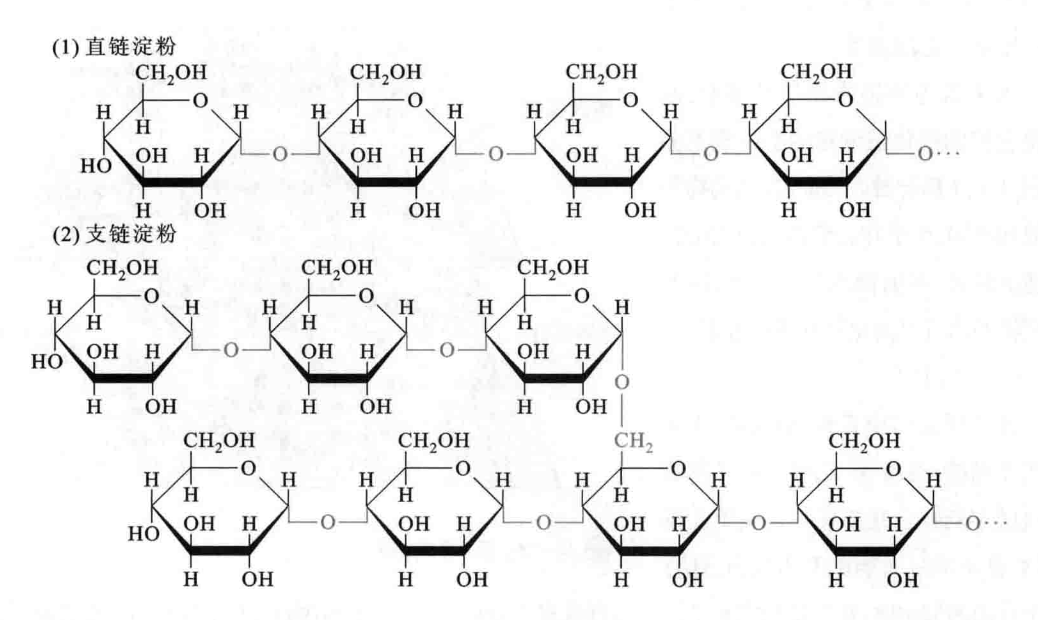


图 15-12　直链淀粉和支链淀粉的化学结构

虽然直链淀粉在水中的溶解性很低，但是它能在水中形成胶束（micelle）悬液。胶束中的多糖链采取螺旋构象。如果这时遇到碘，碘就能插入到疏水螺旋的中间从而呈现蓝色。支链淀粉在水中也能形成胶束结构，但碘与其反应呈现紫红色。

（二）糖原

糖原是动物的贮能多糖，有肝糖原、肌糖原和肾糖原三种形式。其中肝糖原是主要形式，最多能占肝重的 10%。糖原是一种高度分支的分子，在结构上与支链淀粉相似，因此遇到碘显紫红色。但它的分支点更加密集，平均每 8~12 个葡萄糖单位就出现一个分支。糖原分解和合成都是从非还原端展开的，因此，糖原的高度分支结构有利于它的迅速动员和重新合成（参看第二十七章“糖原代谢”），也使得糖原在动物体内特别适合充当一种短期的能源储备。

Quiz6　一个糖原分子具有多少个还原端？

（三）右旋糖酐

右旋糖酐是一种主要以 α-1,6 糖苷键相连的分支多糖，通常存在于酵母和细菌中，其重复的二糖单位主要是异麦芽糖，分支点可能是 α-1→2、1→3 或 1→4 糖苷键。生长在牙齿表面的细菌所产生的右旋糖酐是牙菌斑（dental plaque）的重要成分。细菌产生的右旋糖酐在实验室中经常被用作层析柱的支持介质（如 Sephadex 和 Bio-gel）。

二、结构多糖

最重要的两种结构多糖为纤维素（cellulose）和几丁质（chitin）。两者的结构和性质参看表 15-4。

（一）纤维素

纤维素可以说是世界上最丰富的天然多聚物，几乎所有植物的细胞壁都有它，它对于维持植物细胞的形状十分重要。但纤维素也有其“温柔”的一面，例如，最舒适柔软的天然衣料——棉花几乎就是纯的纤维素。

纤维素与直链淀粉一样，也是一种线性无分支分子，但葡萄糖单位以 β-1,4 糖苷键相连，因此，组成纤维素的二糖单位是纤维二糖。

在 β-1,4 糖苷键的连接方式中，最稳定的构象是葡萄糖单位沿着链交替翻转 180°，因而纤维素链采取的是一种完全伸展的带状构象（extended ribbon）。每条纤维素链除了有链内氢键以外，还与周围的纤维素链形成链间氢键。相邻的几条纤维素链肩并肩地靠在一起，通过链间氢键以及范德华力形成扁平片层。相邻的片层再垛叠在一起，片层内纤维素链呈锯齿状，片层之间的纤维素链也有氢键（图 15-13）。纤维素的上述结构赋予了纤维素一定的强度。

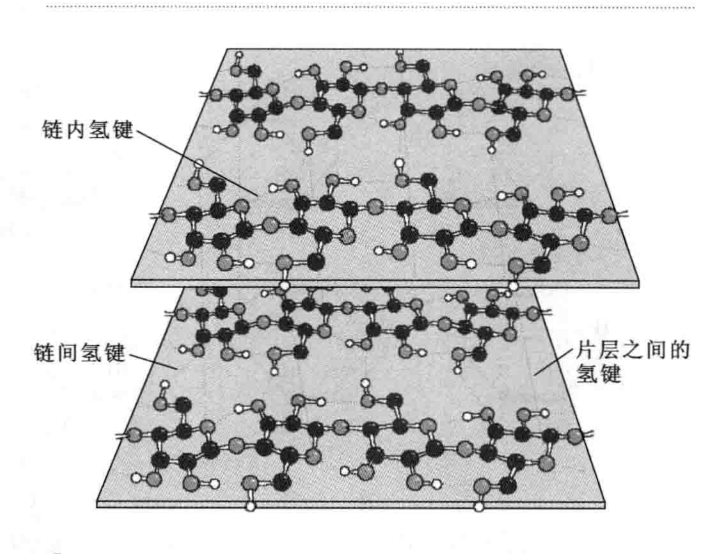


图 15-13 纤维素的构象

大多数动物很难利用纤维素，原因是它们的消化道分泌的水解酶不能识别 β-1,4 糖苷键。然而，反刍动物和白蚁却不同，生活在它们消化道内的细菌能分泌 β- 糖苷酶水解纤维素，因此这两类动物可以消化并利用纤维素。

Quiz7 你能否想出一种对人体无害的灭白蚁的方法？

（二）几丁质

几丁质又名甲壳素、甲壳质、壳聚糖或壳糖胺，是由 *N*- 乙酰葡糖胺形成的直链多聚物。几丁质不仅是甲壳纲动物、昆虫的外骨骼的基本成分，还存在于真菌的细胞壁，在生物功能和结构上与纤维素十分相似。在一级结构上，几丁质由 *N*- 乙酰葡糖胺通过 β-1,4 糖苷键相连，其化学结构可看成 2- 羟基被乙酰胺所取代的纤维素；在高级结构上，它也是伸展的带状构象。但是，纤维素与几丁质在结构上有一个十分重要的差别——片层结构的组织方式。所有片层链的还原端都集中在一侧，非还原端集中在另一侧的排列方式称为平行排列；一个片层内链的还原端与其上下片层内链的还原端位于异侧的排列方式称为反平行排列。天然的纤维素似乎只以平行的方式排列。然而，几丁质有三种形式：全平行排列的 α- 几丁质、全反平行排列的 β- 几丁质和混合型排列的 δ- 几丁质。

Quiz8 纤维素、几丁质、淀粉和糖原都是高度水溶性单糖的多聚物，为什么前两种多糖的水溶性要比后两种差得多？

在自然界中，几丁质资源非常丰富，其含量仅次于纤维素，具有广泛的工业、商业和医用价值。人们已发现，甲壳素有纤维素所不具备的特性，是目前世界上唯一含阳离子的可食性动物纤维，也是除蛋白质、糖、脂肪、维生素、矿物质以外的一种生命要素，因此有人称之为第六生命要素。

e15-5 藻酸、琼脂糖、糖胺聚糖和肽聚糖

除了纤维素和几丁质以外，藻酸（alginates）、琼脂糖、糖胺聚糖和肽聚糖（peptidoglycan）也属于结构多糖，它们也存在于自然界，并具有特殊的生物功能。

框 15-1　生化与健康——蛀牙与防蛀牙

唾液中充满着细菌，浓度高达每毫升唾液含一百万个细菌！在你刷牙后几分钟之内，唾液中的黏蛋白黏附在牙齿表面。然后，立即有数百万计的口腔细菌结合到这个表面，但只有变形链球菌（*Streptococcus mutans*）和乳酸杆菌（*Lactobacillus*）的共同作用才能导致蛀牙。其中变形链球菌能分泌葡聚糖蔗糖酶（dextransucrase），此酶以蔗糖为底物，催化葡糖基的转移，其作用分为两步：第一步是将蔗糖分解成葡萄糖和果糖，第二步是将释放出来的葡萄糖立即转移到一个正在延伸的多糖分子——葡聚糖（右旋糖酐）的一端。葡聚糖像胶水一样，将唾液中的细菌吸附在牙齿的表面，形成牙菌斑。而乳酸杆菌生活在葡聚糖的内部，利用内部的缺氧环境，并以前面从蔗糖分解释放出来的果糖作为原料，进行乳酸发酵，产生的乳酸导致牙齿表面的 pH 降低，使牙釉质钙开始溶解。即使我们每天产生 1 L 唾液，也不能冲走牙齿表面的酸性物质，这是因为葡聚糖对唾液是不透过的。因此养成良好的口腔卫生习惯是很重要的，如每次餐后刷牙和使用牙线，可显著减少牙菌斑的堆积。食用富含钙质的食物，也有助于强化牙齿表面的珐琅质。富含多糖的食品，如水果和蔬菜，也有助于防止蛀牙，原因是葡聚糖蔗糖酶不能使用多糖作为底物，吃水果和蔬菜有助于机械清除牙菌斑。

2010 年 5 月，荷兰格罗宁根大学 Dijkhuizenb 等人得到了葡聚糖蔗糖酶的晶体结构，并阐明了这种酶的催化机制。如果在此基础上，筛选出这种酶的特异性抑制剂，就会为防止蛀牙提供最有效的方法。也许在不远的将来，只要将一点点葡聚糖蔗糖酶的抑制剂添加到牙膏甚至糖果中，龋齿就会成为过去式。

第四节　糖缀合物

糖缀合物（glycoconjugate）是指糖类与非糖物质以共价键相连的复合物。根据非糖物质的本质，糖缀合物可以分为糖脂（参见第十六章“脂质与生物膜”）和糖蛋白两类。

糖蛋白（glycoprotein）和蛋白聚糖（proteoglycan，PG）都是由蛋白质与糖通过共价键相连的复合物，两者都可以看成是蛋白质翻译后发生糖基化修饰的产物。它们的差别在于，前者糖含量占总质量的 1%~60%，由多种具分支的短寡糖链组成；后者糖含量可达总质量的 98%，主要由无分支的糖胺聚糖链组成（典型长度为 80 个单糖单位），糖胺聚糖链多到 100 个以上，有时相对分子质量很大，可达 3×10^6。

与糖胺聚糖共价结合的蛋白质称为核心蛋白，连接方式参见图 15-14。有时，与核心蛋白共价结

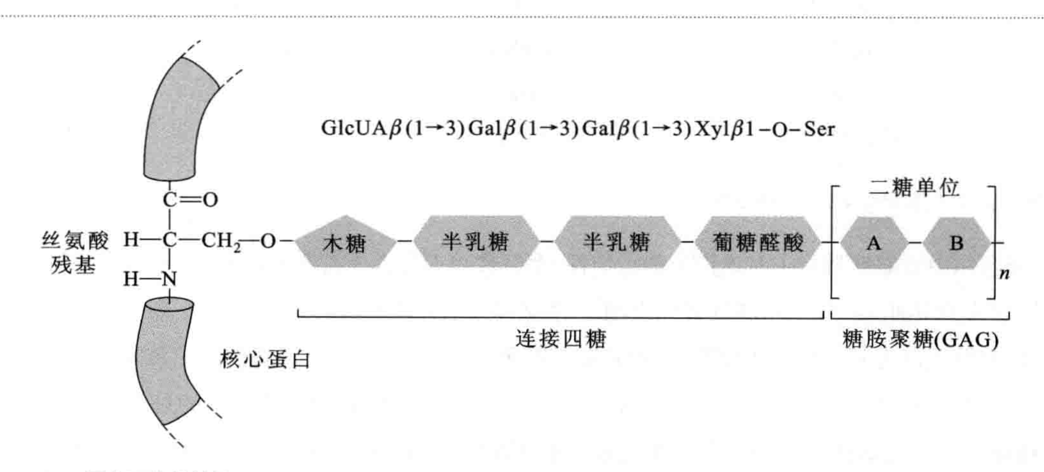


图 15-14　蛋白聚糖的结构

e15-6 六类蛋白聚糖的性质和功能

合的糖胺聚糖不止一种，这样的蛋白聚糖称为多配体蛋白聚糖。已发现至少六类不同的蛋白聚糖，它们在核心蛋白的大小、糖胺聚糖链的数目、糖胺聚糖的类型、体内分布和功能等方面均有差别。

蛋白聚糖的主要功能包括：①调节分泌蛋白的活性；②形成带电的多孔凝胶，限制或促进某些物质的通过；③与某些生长因子结合，调节它们的活性；④在细胞表面作为辅助受体。

糖蛋白是由寡糖链还原端的残基与多肽链特定的氨基酸残基通过糖肽键连接而成，糖肽键主要有 O 型和 N 型（图 15-15）。

O 型糖肽键是由寡糖链还原端残基的半缩醛羟基（一般是 α-GalNAc）与多肽链 Ser/Thr 残基的羟基缩合形成；N 型糖肽键是由寡糖链还原端 β-GlcNAc 残基的半缩醛羟基与多肽链 Asn 的酰胺基缩合而成。后一类糖蛋白的寡糖链由一相同的核心五糖链和若干外围糖基组成（图 15-16）。根据外围糖基的组成特征，其寡糖链可分为高甘露糖型、杂合型和复合型。

(1) O型连接　　(2) N型连接

α-GalNAc　　β-GlcNAc

图 15-15　糖蛋白中寡糖基与蛋白质之间的连接方式

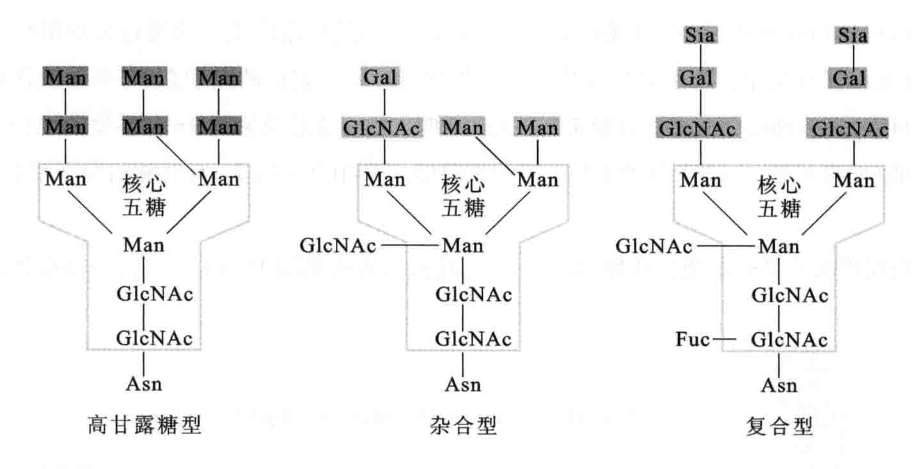


图 15-16　N 型糖蛋白外围寡糖的三种类型

糖蛋白中的寡糖链除了对蛋白质的稳定性、溶解性等理化性质有影响以外，还参与多种生物学功能，例如分子识别、信号转导、多肽链折叠、蛋白质翻译后的分拣和定向等。

许多糖蛋白为分泌蛋白和膜蛋白，例如促红素和卵泡刺激素为分泌性糖蛋白，大多数水溶性激素的受体为整合在质膜上的糖蛋白。如果糖蛋白整合在质膜上，它就会与质膜上的另外一种糖缀化合物（糖脂）一样，其寡糖链部分面向胞外。这些寡糖链在细胞的表面，相当于是一种分子天线，参与细胞与细胞之间的相互作用、细胞识别、细胞黏附以及信号转导。有时，糖蛋白分子上的寡糖链在某些

细胞的表面会形成一个保护性的薄层，如小肠黏膜细胞表面的糖萼可以保护细胞免受蛋白酶的消化。

在许多生物体内，有一类叫凝集素（lectin）的蛋白质可以识别和结合特定的寡糖链。例如，来自刀豆中的凝集素——刀豆素 A（Concanavalin A，Con A），就专门识别非还原端是甘露糖或葡萄糖的寡糖链。有些病毒在感染期间，会使用凝集素附着到宿主细胞表面，例如 H5N1 禽流感病毒株中的 H 表示的就是一种叫血凝素的凝聚素（hemagglutinin）。流感病毒就是通过它识别和结合宿主细胞膜上含有唾液酸的受体糖蛋白而启动感染过程的。

Quiz9 H5N1 中的 N 又代表什么？

推荐网址：

e15-7 本章小结填填看

1. http://themedicalbiochemistrypage.org/carbohydrates.html（完全免费的医学生物化学课程网站，有关糖类部分的内容）
2. http://en.wikipedia.org/wiki/Carbohydrate（维基百科有关糖类的内容）
3. http://chemed.chem.purdue.edu/genchem/topicreview/bp/1biochem/carbo5.html（美国普渡大学一个生化网站，有关糖类部分的内容）
4. http://www2.chemistry.msu.edu/faculty/reusch/virttxtjml/carbhyd.htm（美国密歇根州立大学一个化学网站，有关糖类的内容）

参考文献：

1. Dijkstra. B W，et al. Crystal structure of a 117 kDa glucansucrase fragment provides insight into evolution and product specificity of GH70 enzymes. *Proc Natl Acad Sci*，107：21 406–21 411，2010.
2. Lundt I. Oxidation，reduction and deoxygenation of carbohydrates. *Glycoscience*，1：53–61，2001.
3. Schockman G D，Durst R W. Structure，function and assembly of cell walls of Gram-positive bacteria. *Annu Rev Microbiol*，37：501–527，1983.

第十六章 脂质与生物膜

脂质也称为脂类或类脂，是生物体内一类重要的有机分子，它们所包括的范围很广，其化学结构迥异，生理功能各不相同。然而，脂质都具有一个共同的物理性质，即一般不溶于水或微溶于水，但溶于有机溶剂。许多脂质还有两性的性质，即一个分子同时含有亲水的"头部"和疏水的"尾巴"，这样的性质对于形成生物膜结构至关重要。

本章将主要介绍各种脂质的化学结构和生理功能，并重点讨论两性脂与生物膜的关系。

第一节 脂质的化学结构及其功能

按照化学结构，脂质可分为三大类，即简单脂（simple lipid）、复合脂（compound lipid）和异戊二烯类脂（isoprenoid lipid）。

一、简单脂

图 16-1 脂肪的结构通式

简单脂特指游离的脂肪酸（free fatty acid，FFA）以及由游离的脂肪酸和醇形成的酯，后者包括脂肪（fat）和蜡（wax）。其中脂肪就是三酰甘油（triacylglycerol，TG）或甘油三酯，由 1 分子甘油和 3 分子脂肪酸（fatty acid）通过酯键缩合而成（图 16-1）。生物体内的脂肪酸种类很多（表 16-1），在形成脂肪的时候，参与酯化的 3 个脂肪酸有多种不同的组合，脂肪也因此具有多种形式。

表 16-1 常见脂肪酸的俗称和结构缩写（Δ 表示双键，c 表示顺式）

分类	中文俗称	英文俗称	结构缩写
饱和脂肪酸	月桂酸	lauric acid	12∶0
	豆蔻酸	myristic acid	14∶0
	软脂酸（棕榈酸）	palmitic acid	16∶0
	硬脂酸	stearic acid	18∶0
不饱和脂肪酸	棕榈油酸	palmitoleic acid	16∶1(9) 或 $16:1\Delta^{9c}$
	油酸	oleic acid	18∶1(9) 或 $18:1\Delta^{9c}$
	亚油酸（必需）	linoleic acid	18∶2(9,12) 或 $18:2\Delta^{9c,12c}$
	α- 亚麻酸（必需）	α-linolenic acid（ALA）	18∶3(9,12,15) 或 $18:3\Delta^{9c,12c,15c}$
	γ- 亚麻酸	γ-linolenic acid	18∶3(6,9,12) 或 $18:3\Delta^{6c,9c,12c}$
	花生四烯酸	arachidonic acid（ARA）	20∶4(5,8,11,14) 或 $20:4\Delta^{5c,8c,11c,14c}$
	二十碳五烯酸	eicosapentaenoic acid（EPA）	20∶5(5,8,11,14,17) 或 $20:5\Delta^{5c,8c,11c,14c,17c}$
	二十二碳六烯酸（脑黄金）	docosahexaenoic acid（DHA）	22∶6(4,7,10,13,16,19) 或 $22:6\Delta^{4c,7c,10c,13c,16c,19c}$

按照碳原子数目，脂肪酸可以分为奇数脂肪酸和偶数脂肪酸，但天然的脂肪酸绝大多数为偶数脂肪酸；按照碳链的饱和度，脂肪酸又可以分为饱和与不饱和脂肪酸。例如，豆蔻酸、软脂酸、硬脂酸和月桂酸等为饱和脂肪酸；油酸、亚油酸、亚麻酸和花生四烯酸等为不饱和脂肪酸（表 16-1）。不饱和脂肪酸根据双键的数目又分为单不饱和脂肪酸和多不饱和脂肪酸。天然不饱和脂肪酸的双键多为顺式，某些细菌含有反式的不饱和脂肪酸。不饱和脂肪酸以顺式的形式存在有利于维持膜的流动性。按照营养价值，脂肪酸又分为必需脂肪酸（essential fatty acid，EFA）和非必需脂肪酸（non-essential fatty acid，NEFA），其中必需脂肪酸在人或其他动物体内不能自己合成，但又不可缺少，因此必须从食物中获取。亚油酸和 α- 亚麻酸为必需脂肪酸，其他脂肪酸为非必需脂肪酸。亚油酸是合成花生四烯酸的前体，

而花生四烯酸又是多种前列腺素合成的前体，α- 亚麻酸则是 DHA 和 EPA 的前体。

框 16-1　生化与健康——爱斯基摩人健康的秘密

ALA、DHA 和 EPA（图 16-2）经常被称为 ω-3 脂肪酸，这是因为若从碳链末端的甲基碳（ω- 碳）开始对碳原子进行编号，这三种脂肪酸在 3 号位都有双键。自 20 世纪 70 年代开始，科学家就发现，生活在格陵兰岛的爱斯基摩人因经常食用富含 ω-3 脂肪酸的鱼类而很少患心血管疾病。在此以后，随着人们对 ω-3 脂肪酸的研究逐步深入，发现 ω-3 脂肪酸具有抗炎症、抗血栓和动脉粥样斑形成、降低血脂、舒张血管的功效，经常食用这几种脂肪酸可以显著降低得心血管疾病的风险。EPA 是前列环素合成的前体，而前列环素能抑制血小板凝集，故能减少血栓的形成。DHA 俗称脑黄金，主要存在于动物大脑灰质、精子和视网膜组织中。体内的 DHA 可优先被脑组织吸收，并参入到脑细胞和视网膜细胞膜上的磷脂分子上。脑细胞膜上含有 DHA 的磷脂似乎是轴突形成所必需的。DHA 还能降低血浆中甘油三酯的水平，但其中的机制还不清楚。ALA 能降低心血管疾病的发生率，但究竟是其独自起作用，还是通过 DHA 和 EPA 共同作用，还有待进一步研究。

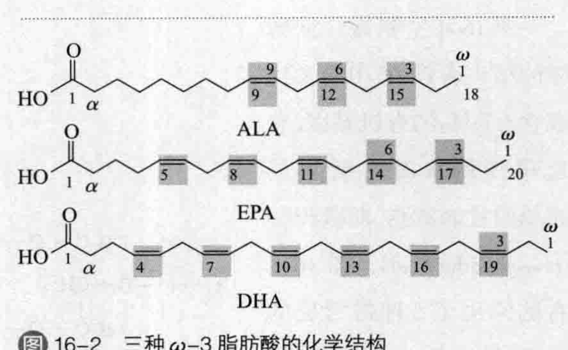


图 16-2　三种 ω-3 脂肪酸的化学结构

鱼类是 ω-3 脂肪酸主要来源食物，含 DHA 和 EPA 的鱼类有：凤尾鱼、鲮鱼、鲱鱼、鲭鱼、鲑鱼、沙丁鱼、鲟鱼、虹鳟鱼和金枪鱼等。植物油；特别是胡桃油、亚麻油和橄榄油，富含 ALA。

脂肪最重要的生理功能是贮存能量，其贮能的效率要高于糖原和蛋白质。例如，1 g 脂肪在体内完全氧化时可释放出 38 kJ 的能量，比 1 g 糖原或 1 g 蛋白质所释放出的能量多两倍以上。此外，脂肪在作为能源贮备的时候，一般贮存在脂肪细胞中，与胆固醇酯一起以脂滴（lipid droplet）的形式存在。脂肪不溶于水，因此不会对细胞的渗透压带来影响。在脂滴的表面通常是一层磷脂单分子层，在外面覆盖一层脂外被蛋白（perilipin）（图 16-3）。这种蛋白质形成了一种保护性屏障，可阻止脂肪酶对脂肪的水解，它的存在有利于脂肪的储存。然而，一旦脂肪细胞受到肾上腺素或胰高血糖素的作用，胞内的蛋白质激酶 A（PKA）很快会被激活。受到 PKA 的催化，脂外被蛋白和脂肪酶均发生磷酸化修饰，磷酸化的脂外被蛋白即丧失活性，导致脂滴表面的保护性屏障开放，而磷酸化的脂肪酶则被激活，并去水解已经暴露在脂滴表面的脂肪分子（参看第二十八章"脂肪、磷脂和糖脂的代谢"）。

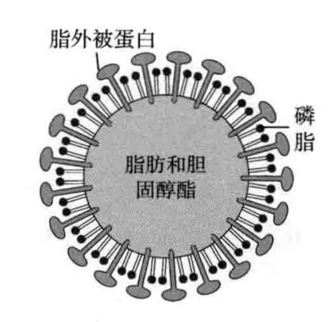


图 16-3　脂滴的结构模型

贮能效率高以及对细胞渗透压没有影响是脂肪作为能源贮备物的优点，但同时它也有两个缺点：一是在机体需要能量的时候，脂肪不能像糖原那样快速地降解，因此更适合作为长期的能源贮备；二是脂肪的氧化分解必须在有氧的条件下才能进行，故在缺氧或无氧的时候，脂肪难以为机体供能。

脂肪除了作为能源贮备以外，还有利于保持体温和保护内脏器官，以及增加水生动物的浮力。例如，生活在南极和北极的动物（如企鹅和北极熊）在皮下都有厚厚的脂肪。某些生物，如骆驼，还能利用脂肪的氧化分解，产生大量的代谢水作为水源。

蜡主要是由长链脂肪酸和高级脂肪醇形成的一种高度不溶于水的酯。它广泛存在于动物的皮毛、植物的叶子和鸟类的羽毛表面，起防水和保护作用。比如，羊毛脂（lanolin）来源于羊的毛皮，是很多护肤品中的有效成分。

二、复合脂

复合脂即复脂，除含有脂酰基和醇基团以外，还含有一些非脂成分，如磷酸基团和糖基。根据非脂成分的不同，复合脂可以进一步分成磷脂（phospholipid）和糖脂（glycolipid）。其中磷脂是含有

磷酸基团的脂质，包括以甘油为骨架的甘油磷脂（glycerophosphatide）和以鞘氨醇为骨架的鞘磷脂（sphingomyelin）。糖脂是含有糖基的脂质。这两大类脂质都属于两性脂，是构成生物膜的主要成分。

（一）甘油磷脂

图 16–4 左侧为甘油磷脂的结构通式，其中的 X 代表含有羟基的有机基团，它是可变的，X=H 时就为最简单的甘油磷脂，即磷脂酸（phosphatidic acid）。图 16–4 右侧给出了 5 种最常见的甘油磷脂的 X 基团，这 5 种甘油磷脂是磷脂酰胆碱（phosphatidylcholine，PC）、磷脂酰肌醇（phosphatidylinositol，PI）、磷脂酰丝氨酸（phosphatidylserine，PS）、磷脂酰乙醇胺（phosphatidylethalamine，PE）和二磷脂酰基甘油（diphosphatidylglycerol），它们都可以视为磷脂酸的衍生物。其中，磷脂酰胆碱又称为卵磷脂（lecithin），二磷脂酰甘油又称为心磷脂。

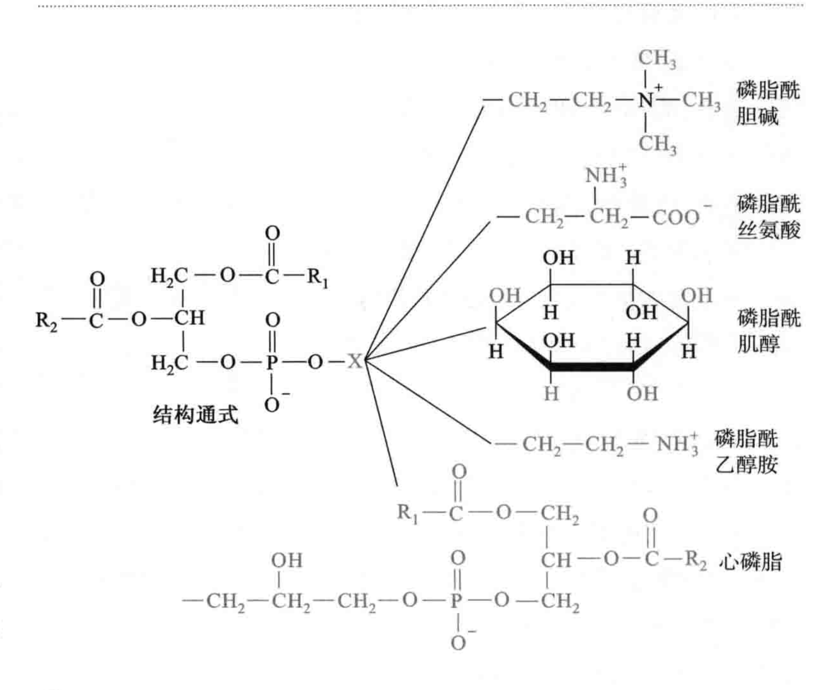


图 16–4　甘油磷脂的结构通式及其五种常见磷脂的化学结构

Quiz1 计算一下在生理 pH 下，各种磷脂带电状况如何？

以磷脂酰胆碱为例，整个甘油磷脂分子可分成两个部分（图 16–5）：一部分由两个长长的非极性的碳氢链构成，具有疏水的性质，有人形象地称之为疏水尾巴；另一部分由极性的磷酸化的 X 基团组成，具有亲水的性质，被称为亲水头部。甘油磷脂的这种特殊结构使它成为一种典型的两性分子，特别适合作为生物膜骨架，因此广泛存在于各种膜中。

Quiz2 你认为甘油磷脂有没有等电点？

甘油磷脂中有一种特殊的磷脂，称为醚磷脂（ether glycerophospholipid），其 1 号位和 / 或 2 号位的脂酰基变成了脂醚基。古菌绝大多数生活在恶劣或极端的环境之中，如强嗜热古菌（*Hyperthermophilic archaea*），为了适应环境，其细胞膜主要由醚磷脂组成，原因是醚键比酯键更加稳定。此外，古菌的醚磷脂构型主要为 S 型，而不是细菌和真核生物的 R 型。

已发现两种类型的醚磷脂（图 16–6）：一类为甘油二醚（glycerol diether），其脂醚基为含有分支的长链碳氢链（20 碳）；另一类为二甘油四醚（diglycerol tetraether），2 个含有分支的长链碳氢链（40 碳）在两端各与 1 个甘油分子以醚键相连。

（二）鞘磷脂

鞘磷脂（sphingomyeline）的结构与甘油磷脂十分相似，也是一种两性分子，只是由神经鞘氨醇（sphingosine）代替了甘油。图 16–7 显示了神经鞘氨醇的分子结构，当它的氨基被脂酰化以后，形成的化合物就是神经酰胺（ceramide）。图 16–8 还显示了鞘磷脂的结构通式，X 基团通常为胆碱或乙醇胺。

鞘磷脂也是生物膜的重要组分，不过在动物的神经组织中的含量较高，特别是在髓鞘。

（三）糖脂

糖脂是糖类通过它的半缩醛羟基与脂质以糖苷键连接而成的糖缀化合物。包括鞘糖脂（glycosphingolipid）和甘油糖脂（glyceroglycolipid），其中鞘糖脂的脂质部分是神经鞘氨醇，而甘油糖脂的脂质部分为二酰甘油（diacylglycerol，DG）即甘油二酯。鞘糖脂和鞘磷脂则通称为鞘脂（sphingolipid）。

鞘糖脂又分为中性鞘糖脂和酸性鞘糖脂。前者的糖基无唾液酸成分，通常为单糖、双糖、三糖或其他寡糖。半乳糖神经酰胺是第一种得到鉴定的鞘糖脂（图 16–8），因最早发现于人脑，又名为脑

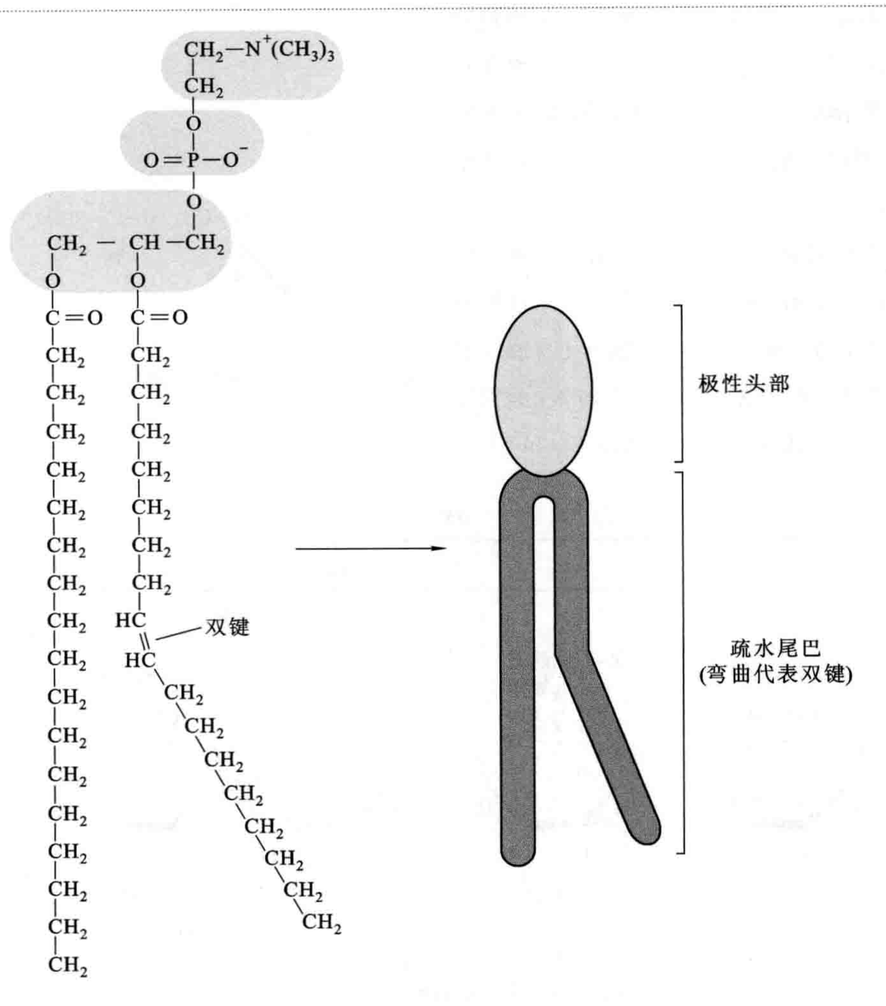


图 16-5 甘油磷脂的两性性质

(1) 甘油二醚

醚键

C20

(2) 二甘油四醚

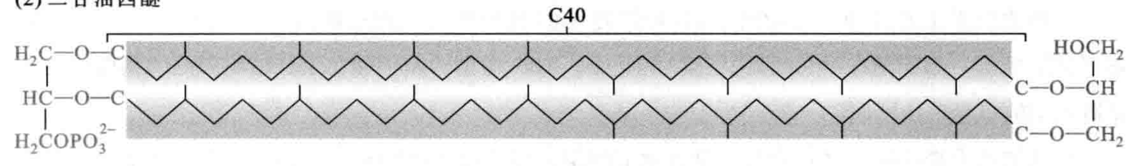


图 16-6 甘油二醚和二甘油四醚的化学结构

神经鞘氨醇

神经酰胺

鞘磷脂

图 16-7 神经酰胺和神经鞘氨醇的化学结构及鞘磷脂的结构通式

苷脂(cerebroside);后者的糖基含有酸性的硫酸化糖基或唾液酸。图 16-9 显示的是一种含有唾液酸的鞘糖脂,其糖基是一条寡糖链,由于它广泛存在于动物的神经节,又名为神经节苷脂(ganglioside)。

神经节苷脂的命名很特别,以 G 代表神经节苷脂,M(mono)、D(di)和 T(tri)分别表示寡糖链上含有 1、2 和 3 个唾液酸,下标数字 1、2 和 3 表示与唾液酸相连的寡糖链序列(图 16-9),分别是 Gal-GalNAc-Gal-Glc、GalNAc-Gal-Glc 和 Gal-Glc。

图 16-8 中性鞘糖脂的化学结构

图 16-9 酸性鞘糖脂的化学结构

鞘糖脂在膜上参与细胞之间的通信,并作为 ABO 血型的抗原决定簇,有些鞘糖脂还作为一些病毒(如仙台病毒)和细菌毒素(如霍乱毒素)的受体。ABO 血型的差别就在于红细胞表面的鞘糖脂在寡糖链末端的单糖单位不同。

甘油糖脂由糖基通过糖苷键与二酰甘油上的游离羟基相连(图 16-10),其中单半乳糖基二酰甘油和二半乳糖基二酰甘油最常见,它们主要存在于植物的叶绿体膜和微生物的细胞膜。

图 16-10 甘油糖脂的化学结构

三、异戊二烯类脂

此类脂衍生于异戊二烯，在结构上可被剖析成若干个异戊二烯单位，主要包括萜（terpene）、脂溶性维生素和类固醇。虽然它们在体内的含量并不高，但却具有非常重要的生物学功能。

（一）萜

就分子结构而言，萜是由若干个异戊二烯单位连接而成，连接的方式主要是“头尾”相连（head to tail），也有“尾尾相连”（tail to tail）（图 16–11）。

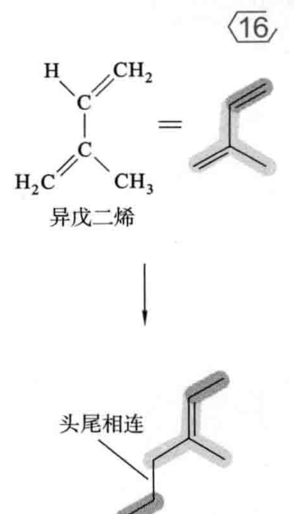


图 16–11　萜的形成

根据所含的异戊二烯单位的数目，萜可分为单萜（monoterpene，2 个异戊二烯）、倍半萜（sesquiterpene，3 个异戊二烯）、双萜（diterpene，4 个异戊二烯）、三萜（triterpene，6 个异戊二烯）、四萜（tetraterpene，8 个异戊二烯）和多萜（polyterpene，多个异戊二烯）。特别值得一提的是番茄红素，它与 β- 胡萝卜素的结构相似，为西红柿红色色素的主要成分。已有很多证据表明，番茄红素在体内可作为一种出色的抗氧化剂，能防止自由基对机体的损害，具有抗衰老、抗癌和抗心血管病的功效。

（二）类固醇

类固醇也称甾类（steroid），其核心结构是由 3 个六元环和 1 个五元环融合而成的环戊烷多氢菲（perhydrocyclopentanophenanthrene）。胆固醇（cholesterol）是最重要和最常见的类固醇，其他甾类几乎都由它衍生而来。

1. 胆固醇及其衍生物

胆固醇的分子结构如图 16–12 所示，整个分子只有位于 3 号位的羟基是亲水的，其余部分完全由疏水的碳氢链组成，因此胆固醇是一种疏水性较强的两性脂。

图 16–12　胆固醇的结构

胆固醇的两性性质使得它能够与磷脂和糖脂一起构成膜，但在膜上的功能主要是调节膜的流动性。胆固醇还是体内许多重要活性物质的前体，例如维生素 D、固醇类激素和胆汁酸（bile acid）都是以它作为原料合成的。

2. 其他固醇

胆固醇主要存在于动物体内，在植物和某些微生物体内仅有它的类似物，这些类似物的结构与胆固醇十分相似。例如，豆科植物中普遍存在的豆固醇（stigmasterol）、谷物类植物中发现的谷固醇（sitosterol）和某些真菌内发现的麦角固醇（ergosterol）。其中麦角固醇被动物吸收后，在体内经紫外线的照射，可转化为维生素 D_2 的前体。

（三）脂溶性维生素

维生素 A、D、E 和 K 均由异戊二烯单位聚合而来，不溶于水，其结构与功能参见第十四章“维生素与辅酶”。

第二节　生物膜的结构及其功能

生物膜是指由糖类、两性脂类和蛋白质组装而成的一种薄膜状结构，厚度为 7~10 nm，包括细胞质膜和真核细胞所特有的细胞器膜（即内膜系统）。在细胞中，生物膜能行使多种功能。首先，它充当一

种选择透过性屏障，将细胞内外环境有效地隔离开来，以维持两种环境在组分上的差异，为细胞的生命活动提供相对稳定的内环境，或者在真核细胞内将不同的细胞器隔离开来，既提高了各种细胞活动的效率，又方便了对它们的调控。其次，生物膜还参与物质转运、能量转换、神经传导和信号转导等一系列重要的生理过程。许多疾病的发生也与生物膜有直接的关系。

一、生物膜的化学组成

从化学组成上来看，生物膜主要是由两性脂和蛋白质组装成的超分子复合物，某些生物膜还含有糖类。膜上各成分的比例随着膜的来源和功能上的不同而有所差异。在绝大多数的膜上，蛋白质和脂质含量的比值≥1。一个显著的例外是髓鞘，它位于许多神经纤维上，充当生物电的绝缘体。在某些膜上，蛋白质与脂质的比例高达4:1；而在另外一些膜上，两者的比例只有1:4。一般而言，膜的功能越复杂，其蛋白质含量就越高。在膜上，游离的糖分子极为少见，它们通常与脂质和蛋白质以共价键结合而构成糖脂和糖蛋白。

Quiz3 对于线粒体内膜和外膜而言，你认为哪一种膜上的蛋白质含量更多？为什么？

（一）膜脂

构成生物膜的脂主要包括磷脂、糖脂和胆固醇三类，它们都是两性脂，其中磷脂的含量最高。胆固醇主要存在于动物细胞的质膜以及除了线粒体内膜以外的细胞器膜上。原核细胞的质膜上一般没有胆固醇，支原体是一个例外。植物细胞的膜上含有大量的植物固醇（phytosterol），如菜油固醇（campesterol）、谷固醇和豆固醇，只有少量的胆固醇。真菌的膜上无胆固醇，但含有结构与胆固醇相似的麦角固醇。某些抗真菌的药物，如制霉菌素（nystatin）和两性霉素 B（amphotericin B）之所以能够杀死真菌而对人体基本无害，是因为它们能够特异性地与真菌细胞膜上的麦角固醇结合，然后在膜上形成小孔，导致 K^+ 等发生泄漏。

（二）膜蛋白

膜蛋白是生物膜的又一重要成分，与生物膜的功能直接相关，包括定位于膜上的酶、受体、离子通道、转运蛋白、结构蛋白、抗原和电子传递体等。这些蛋白质在结构上差异很大，因此很难用一个通用的模型来概括它们的结构特征，但根据它们在膜上的性质，可分成外周蛋白（peripheral protein）、内在蛋白（intrinsic protein）和脂锚定蛋白（lipid-anchored protein）（图 16-13）。

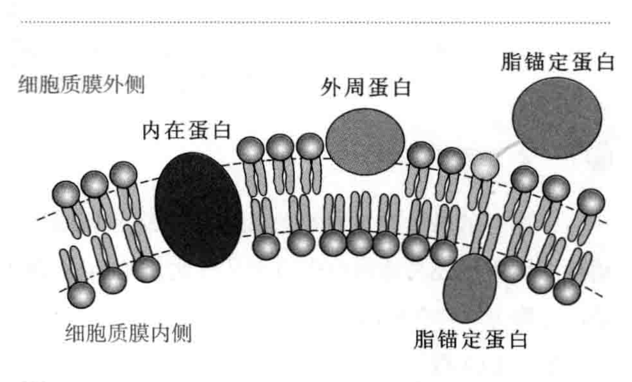


图 16-13 膜蛋白的分类（以细胞质膜为例）

1. 外周蛋白

也称作外在蛋白（extrinsic protein），分布在膜的表面，在三维结构上与球状蛋白质很相似，有的通过离子键或氢键与膜表面的亲水头部，或者与膜内在蛋白暴露在外的部分松散结合，有的则通过自身的一个疏水的小环插入到膜内。一般使用高盐或高 pH 溶液很容易将它们从膜上除去。例如，细胞色素 c 就是线粒体内膜上的一种外周蛋白（参看第二十章“生物氧化”）。

Quiz4 为什么高盐或高 pH 溶液可将外周蛋白从膜上分离？

2. 内在蛋白

也称整合蛋白（integral protein），与膜脂一样，也属于两性分子，但与膜脂不同的是，其亲水区和疏水区主要分别由亲水氨基酸残基和疏水氨基酸残基的 R 基团提供。在生物膜上，内在蛋白的疏水区域与膜内的疏水核心通过疏水键紧密结合，而亲水区则伸展在水相之中，或者在膜内形成亲水性通道。内在蛋白与膜脂的这种紧密结合，使得它们很难与生物膜分离，只有使用对膜具破坏性的试剂（如有机溶剂或去污剂）才能将它们与膜脂分开。此外，绝大多数内在蛋白含有一个或几个跨膜的肽段（如与 G 蛋白偶联的受体蛋白具有 7 次跨膜的结构）。由于 α 螺旋能最大限度地降低肽键本身的亲水性

质，且在疏水环境中更加稳定，因此这些跨膜的肽段主要是由疏水的 α 螺旋构成。但也有蛋白质通过 β 桶结构跨膜（参看第二章“蛋白质的结构”）。

3. 脂锚定蛋白

这类膜蛋白通过共价相连的脂质固定到质膜上，虽然并不含有跨膜的肽段，但通过共价结合的脂质与膜牢固地结合。它们仿佛是停泊在岸边的小船，河岸就像细胞质膜，锚就像脂质，船就像是蛋白质，而河流就像是细胞质基质。脂锚定蛋白有 4 类：

(1) 第一类通过 N 端 Gly 残基的氨基与脂肪酸（主要是豆蔻酸或软脂酸）的羧基形成酰胺键，然后通过疏水的脂酰基与膜脂疏水核心的结合而锚定在细胞质膜上。

(2) 第二类通过硫酯键（Cys 的巯基参与）或酯键（Ser 和 Thr 的羟基参与）与脂酰基相连，然后同样通过疏水的脂酰基与膜脂疏水核心的结合，而锚定在细胞质膜上（图 16–14）。

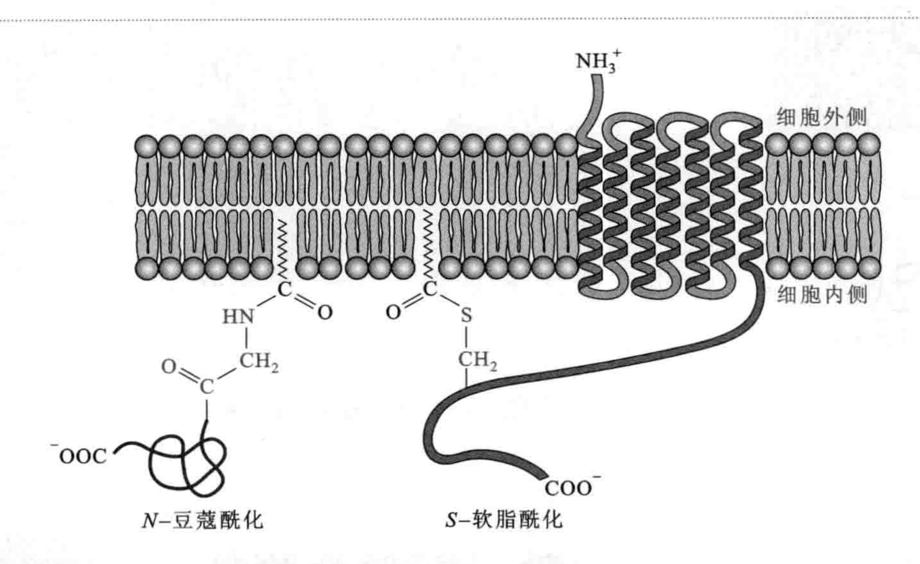


图 16–14　通过酰胺键或酯键相连的脂锚定蛋白

(3) 第三类是通过硫醚键相连的异戊二烯化蛋白（prenylated protein）。这类脂锚定蛋白含有若干个与 Cys 残基的巯基共价相连的异戊二烯基团，通过疏水性的异戊二烯基团插入到脂双层的内部而锚定在膜上（图 16–15）。异戊二烯基团是通过法尼基化（farnesylation）或牻牛儿牻牛儿基化（geranylgeranylation）反应引入的（参看第三十章“胆固醇代谢”）。异戊二烯基团总是连接在 CAAX 序列之中的 Cys 残基上（C=Cys，A=aliphatic（脂肪族氨基酸残基），X= 任何残基）。

参与受体酪氨酸激酶信号转导通路的关键成分——Ras 蛋白，就是通过这种方式锚定在细胞质膜内侧的。已发现约 1/3 的癌症与 Ras 蛋白的突变有关联，因此科学家在考虑使用催化 Ras 蛋白发生法尼基化的法尼基转移酶的抑制剂来治疗癌症。

(4) 最后一类称为糖基磷脂酰肌醇（glycosylphosphatidyl inositol，GPI）锚定蛋白，其 C 端的氨基酸残基与磷酸化的乙醇胺相连，后者再与膜上糖基化的磷脂酰肌醇共价相连（图 16–16）。这一类脂锚定蛋白主要存在于动物细胞。许多动物细胞表面的抗原和黏附分子通过这种方式固定到质膜上。

二、生物膜的基本结构与性质

生物膜的基本结构是由膜脂和膜蛋白的基本性质决定的，其最基本的结构骨架是双层的膜脂分子，简称脂双层结构（图 16–17）。

（一）脂双层结构

膜脂的两性性质使得它们在水相中能自发地形成 4 种在热力学上稳定的结构，即平面脂单层、胶束微粒（micelle）、平面脂双层和封闭的脂双层结构（图 16–18）。在这几种结构中，膜脂分子的疏水尾部通过疏水键和范德华力紧密结合在一起，以尽量避免与水分子接触，而亲水的头部则暴露在水相之

Quiz5 如果将磷脂分子扔到脂溶性溶剂（如苯和氯仿）中，那会形成什么样的膜结构？

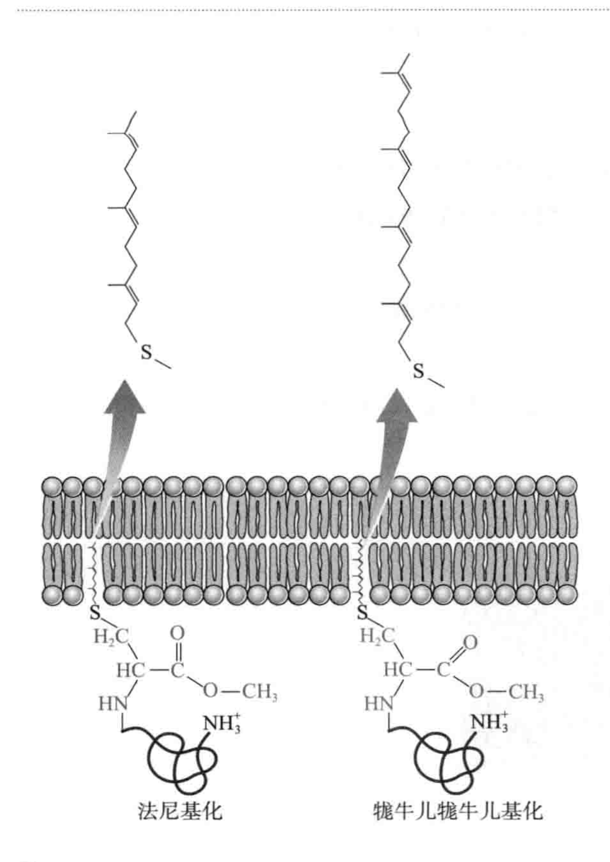


图 16-15　异戊二烯化脂锚定蛋白



图 16-16　GPI 锚定蛋白的结构

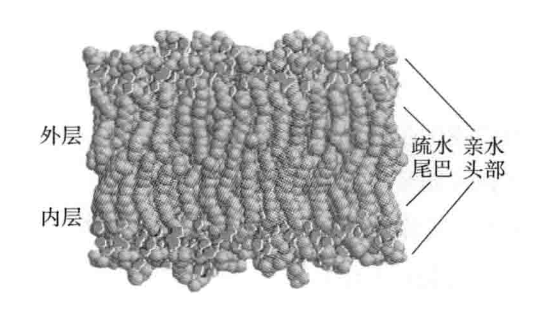


图 16-17　脂双层结构的空间填充模型

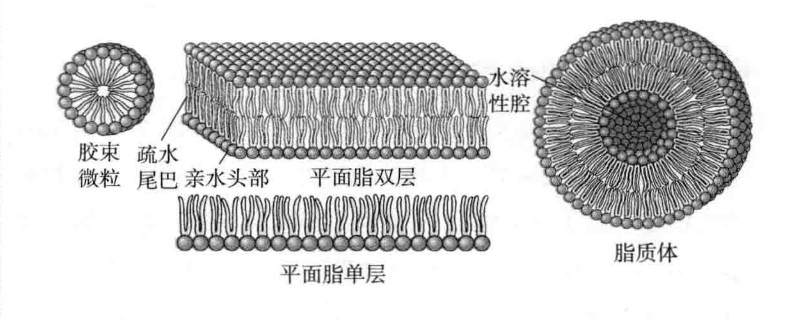
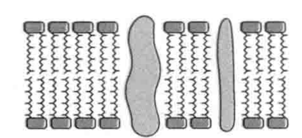
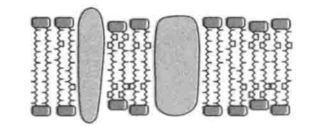


图 16-18　膜脂在水相中自组装形成的几种结构

中。其中封闭型的脂双层结构称为脂质体(liposome),这种结构将疏水尾巴隐藏得最好,因此是最稳定的。1925 年由荷兰科学家 E. Gorter 和 F. Grendel 最早提出的生物膜结构模型即是这种脂双层结构,该模型可以解释生物膜许多重要的现象。



带有内在蛋白的双层脂膜



带有内在蛋白的单层脂膜

图 16-19　强嗜热古菌的膜结构图

强嗜热古菌的细胞膜在某些区域为脂单层结构,构成这一部分的磷脂为二甘油四醚,其厚度相当于两层的甘油二醚(图 16-19)。

(二) 生物膜的结构模型

目前最能反映生物膜化学结构和生物学功能的结构模型是流动镶嵌模型(fluid mosaic model)(图 16-20),它由 S. Jonathan Singer 和 Garth Nicholson 于 1972 年提出。该模型的主要内容是:生物膜以具有流动性的二维脂双层作为基本骨架,在上面镶嵌着各种蛋白

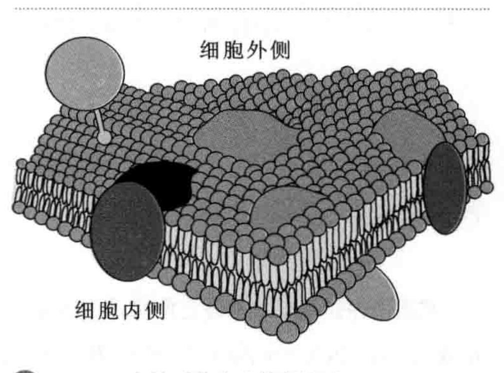


图 16-20　生物膜的流动镶嵌模型

质，这些蛋白质有的分布在表面，有的横跨整个脂双层，相对自由地“浮动”在脂双层骨架上。在这里，可以将脂双层比作海洋，膜蛋白比作冰山，膜脂与膜蛋白的关系就像是海洋和冰山。

框 16-2　生化动态——脂质体的应用

脂质体是一种由磷脂双分子层包被的囊泡状结构，最早由 Alec Bangham 于 1961 年发现。当他将磷脂和水混合以后，施加超声波处理，脂质体便会自发地形成。自那以后，脂质体不但作为一种简单的膜系统，用于研究膜的各种功能，而且已成为运转多种药物的高效输送系统——从抗肿瘤和抗病毒药物，到治疗脱发的药物。

如果将药物与磷脂溶液放在一起，脂质体在形成过程中就可以把药物包被在里面。这些脂质体可静脉注射或直接用于身体表面。若在脂质体表面引入亲水性分子，可增加它们在血流中存留的时间。为了增加脂质体作用的特异性，还可在脂质体的表面引入导向性分子，如对特定癌细胞特异性的抗体。这样的脂质体引入到体内，只会定向地将药物输送到特定细胞，可避免其误伤正常健康的细胞。脂质体还可应用于化妆品，将维生素、中草药剂或其他营养元素包装到脂质体中。当用到皮肤上时，脂质体很容易通过外层的死皮，与内层活的皮肤细胞融合，从而将有效成分输送到皮肤细胞内。

流动镶嵌模型虽然有许多实验证据支持，但它描绘的生物膜仍然是一种相对静止的结构，而生物膜真实的结构更为复杂和更具动态。

(三) 生物膜的基本性质

通过对各种天然生物膜和人工膜的研究，发现生物膜至少具有以下几个基本的性质：

1. 不对称性

生物膜的不对称性包含了结构的不对称性和功能的不对称性，前者又分为膜脂的不对称性和膜蛋白的不对称性，而正是结构的不对称性决定了功能的不对称性。

膜脂的不对称性是指不同的膜脂在脂双层内外两层中的差异分布。例如，含有胆碱的磷脂（磷脂酰胆碱和鞘磷脂）主要分布在脂双层的外层，它们在生理 pH 下不带有净电荷，因而不会相互排斥。而氨基磷脂（磷脂酰丝氨酸和磷脂酰乙醇胺）主要分布在内层，由于磷脂酰丝氨酸带有净负电荷，从而使得质膜的内面带负电荷。然而，细胞一旦发生凋亡，磷脂酰丝氨酸就会开始从膜的内层迁移到外层。磷脂酰丝氨酸的外移也因此经常作为检测细胞发生凋亡的早期信号。再如，在动物细胞膜上，外层的胆固醇含量高于内层。关于膜脂不对称性形成的机理尚不十分清楚，但已发现一种叫翻转酶（flippase）的蛋白质可能在其中起一定作用。如图 16-21 所示，当一个脂分子扩散到翻转酶的时候，翻转酶可催化它转移到膜的另一层。有两类翻转酶：一类被动地作用，不需要能量；另一类主动作用，需要 ATP 水解提供能量。显然，只有起主动作用的翻转酶才可导致膜脂的不对称性。

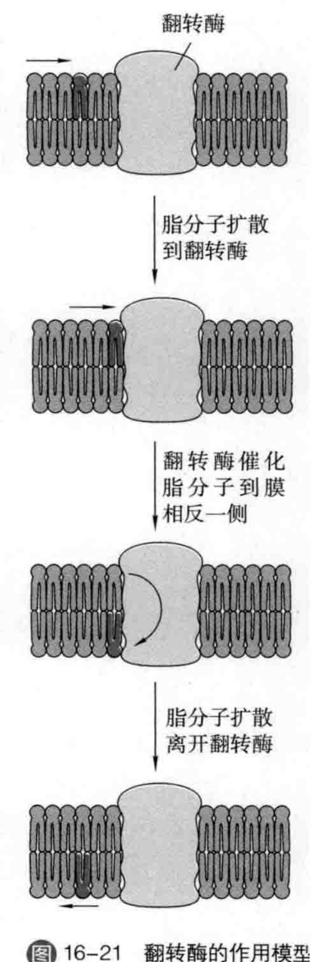


图 16-21　翻转酶的作用模型

膜脂的不对称性在功能上十分重要，特别是对于细胞的某些信号通路。例如，作为第二信使（二酰甘油和 1,4,5- 三磷酸肌醇）前体的磷脂酰肌醇主要位于质膜内层（参看第十七章“激素及其受体介导的信号转导”）。

膜蛋白在脂双层上分布的不对称性主要反映在：有的外周蛋白只附着在外层的表面，而有的正好相反；脂锚定蛋白与外周蛋白的情形差不多；内在蛋白也一样，有的插在外层，有的只插在内层，有的虽然横跨内外两层，但是相连的寡糖链只存在于膜的外侧。膜蛋白分布的不对称性对于生物膜的正常生理功能是极为重要的。例如，跨膜转运中，质膜受体上与配体结合的部位面向胞外才有效。至于膜蛋白在膜上不对称性的建立，与膜蛋白在细胞内合成以后所带有的信号序列以及胞内的定向、分拣机制有关（参看第四十章“蛋白质的翻译后加工及其定向和分拣”）。

2. 流动性

生物膜的流动性是指生物膜各组分所作的各种形式的运动，其中膜脂的流动性（fluidity）更强。

膜脂分子能做的运动包括侧向扩散、转动、头部的伸缩和上下翻转等(图 16-22)。其中上下翻转需要磷脂分子的亲水头部穿过疏水核心,这在热力学上是不利的,因此发生的频率较低。

膜脂的流动性与膜所处的温度、脂酰基的碳链长度和不饱和性,以及胆固醇的含量有关。温度越高,膜的流动性越强。饱和的碳氢链使疏水尾巴能更紧密有序地聚集在一起,从而导致范德华力增强。较长的碳氢链有类似的效果。而不饱和的碳氢链含有顺式的双键,这种双键将剧烈的弯曲引入到疏水尾巴上,不利于疏水尾巴之间紧密有序的聚集。因此,脂酰基的碳链越短、不饱和度越高,膜的流动性就越强。

然而,胆固醇作为一种在结构上十分特别的膜脂,对生物膜的流动性有双面影响(图 16-23):在膜的流动性较强的时候,其特殊的环结构使其能够“见缝插针”,填充在不饱和碳氢链因弯曲产生的缝隙之中,像胶水一样,加强疏水尾巴之间的范德华力,这时候的胆固醇使得膜的流动性降低;相反,如果膜的流动性不高,其刚性的环结构就夹杂在饱和的碳氢链之间,不利于疏水尾巴之间紧密有序的聚集,这时候它反而能够提高膜的流动性。正因为如此,有人将胆固醇称为动物细胞膜流动性的缓冲剂。

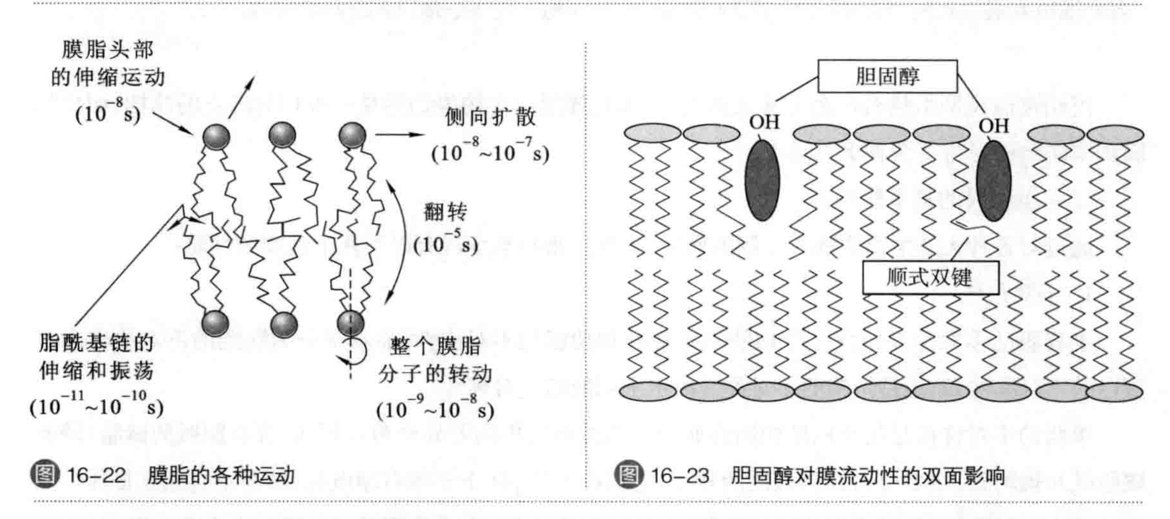


图 16-22 膜脂的各种运动

图 16-23 胆固醇对膜流动性的双面影响

虽然绝大多数原核细胞和植物细胞的细胞膜上没有胆固醇,但它们能够通过改变膜脂分子中不饱和脂肪酸和短链脂肪酸的比例来调节膜的流动性。例如,温度的下降让微球菌(*Micrococcus*)用更多的十六碳脂肪酸取代原来磷脂分子中的十八碳脂肪酸。再如,降低大肠杆菌的培养温度,就很快发现其膜上的不饱和脂肪酸的比例开始上升。在环境温度下降的情况下,大肠杆菌质膜上不饱和脂肪酸的含量上升,是因为温度的下降诱导了其胞内负责在脂肪酸分子上引入双键的去饱和酶(desaturase)的基因表达。与大肠杆菌相似,酵母和植物也是通过改变膜上不饱和脂肪酸的比例来抵抗温度变化对膜流动性的影响的。但差别在于,温度下降并不能提高它们体内的去饱和酶基因的表达,而是增加了去饱和酶的底物之一——氧气的溶解性,从而使去饱和酶的活性增强。

因为有胆固醇对膜流动性的缓冲作用,且高等动物体温恒定,因此绝大多数高等动物细胞质膜上脂肪酸碳链的长度和不饱和度相对稳定。不过,驯鹿这种动物非常有趣!有研究表明,越靠近蹄子部位的细胞,越接近冰冷的雪地,温度就越低,膜上的不饱和脂肪酸和短链脂肪酸的含量就越高!

膜蛋白的流动性主要有两种形式:一种是沿着与膜平面垂直的轴做旋转运动,如位于嗜盐菌质膜上的菌紫质;另一种是沿着膜表面做侧向扩散运动,这是绝大多数膜蛋白运动采取的方式。这种侧向扩散在多数情况下是随意和无序的。然而,某些膜蛋白并非随意扩散,而是集中到细胞膜的某一固定区域,例如,许多受体蛋白在结合配体以后,会聚合在细胞膜某一特定区域。

流动性对生物膜的功能具有深刻的影响。比如,随着膜的流动性增强,膜对水和其他亲水性小分子的通透性就会增加。再如,许多生长因子在与膜上的受体结合以后,能够提高质膜在局部的流动性,从而有助于受体之间形成二聚体,而受体的二聚化对于随后的信号转导至关重要。

3. 生物膜的相变和分相

相变和分相与膜的流动性有密切的关系，是生物膜结构的又一重要特征。在比较低的温度下，膜脂分子上的脂酰基链呈高度有序的排列，这种有序排列导致生物膜处于一种类似固体或晶体的状态。但随着温度的升高，膜脂上的脂酰基从有序状变为无序，原先处于固态的膜脂逐渐转变为流体态或液晶态。生物膜这种从固态转变为液态的过程被称为相变。发生相变时的温度被称为相变温度（transition temperature）。脂酰基链越长、饱和度越高，生物膜的相变温度就越高。在生理温度下，膜脂双分子层中相当一部分表现为流体态或液晶态，而另一部分则表现为固体态或结晶态，生物膜的这种性质被称为分相（图 16–24）。

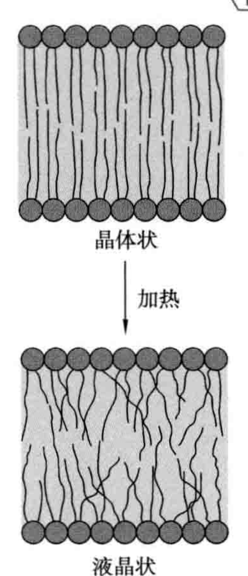


图 16–24　膜脂的相变

膜脂的相变有可能影响到膜蛋白的功能。如果一种内在蛋白的活性中心位于它的亲水区域，相变对该蛋白质的功能影响就不大。但是，如果这种内在蛋白是一种跨膜的转运蛋白或受体蛋白，膜脂的相变就可直接影响到它的转运速率或与激素之间的亲和性。胰岛素受体就是一例，随着膜脂中不饱和脂肪酸比例的增加（通过将细胞放在富含不饱和脂肪酸的培养基上进行培养而实现），受体与胰岛素的亲和力逐步增大。

同一个膜的不同区域的相变温度可因为局部的膜脂成分不同而有所不同。例如，有一种叫脂筏（lipid raft）的结构，是质膜上富含胆固醇和鞘磷脂的微结构域（microdomain）。这种膜结构的相变温度要高于膜的其他部分，其存在具有特殊的生物学功能。

e16–1 脂筏概况

第三节　物质的跨膜转运

细胞要生存和繁殖，就需要及时、不断地从周围环境中得到营养物质，与此同时还需要及时清除细胞内的代谢废物。可是，所有物质进出细胞或细胞器都需要通过生物膜屏障，那么物质是怎样进出细胞的呢？

对于离子与小分子物质而言，物质的跨膜转运可分为两种形式：一种是被动转运（passive transport），另一种是主动转运（active transport）。两者的差别一方面是转运的方向，是逆浓度梯度还是顺浓度梯度，另外一方面是表现在转运过程中对能量的依赖性。在被动转运的时候，被转运的物质从浓度高的一侧向浓度低的一侧转移，此过程在热力学上是有利的，不需要消耗能量；而主动转运则是逆浓度梯度进行的，需要提供能量，以克服能障。存在于细胞膜上的各种转运系统参见表 16–2 和图 16–25。

Quiz6 你认为葡萄糖在肾小管的重吸收是以什么机制进行的？为什么高血糖会导致糖尿？

表 16–2　跨膜转运的几种方式的比较

转运方式	载体蛋白	饱和动力学	转运方向	能量消耗
简单扩散	无	无	顺浓度梯度	无
通道或孔	有	无	顺浓度梯度	无
易化扩散	有	是	顺浓度梯度	无
初级主动转运	有	是	逆浓度梯度	直接消耗，通常是 ATP 的水解
次级主动转运	有	是	逆浓度梯度	间接消耗，为离子梯度

被动转运又分为简单扩散（simple diffusion）和易化扩散（facilitated diffusion），两者的差别在于后者需要蛋白质的主动参与。非极性的小分子和不带电荷的极性小分子，例如各种气体（O_2、N_2、CH_4 和 NO 等）、脂溶性激素、脂溶性维生素和乙醇，很容易通过简单扩散的方式进出细胞或细胞器，而离子和较大的极性分子，如 Na^+ 和葡萄糖，需要在由膜蛋白形成的通道（channel）和转运蛋白（transporter）的帮助下，通过通道运输或易化扩散才能出入细胞或细胞器。如果一种分子或离子通过由某些蛋白质形成的跨膜通道或孔进行转运，这就是通道运输。通道运输与简单扩散十分相似。例如水通过单纯的脂双层膜是比较慢的，但水孔蛋白（水通道蛋白；aquaporin，AQP）则让水分子以更快的速率（每秒钟

e16–2 简单扩散和易化扩散

10^9~10^{11} 个水分子通过)进出细胞。有的通道一直处于开放状态,如钾泄漏通道;有的通道则受到严格的调控,平时处于关闭状态,仅在细胞需要的时候才会"开闸放流"。这些通道有的需要特殊的配体(如乙酰胆碱和 cGMP)结合才能开放,有的受膜电位的控制,有的受机械力(摩擦力、压力、牵拉力、重力或剪切力等)控制,有的受磷酸化控制(参看第十七章"激素及其受体介导的信号转导")。例如,植物细胞的水通道蛋白比动物多了一个肽环,在干旱和水涝的时候,这个肽环像一个塞子一样,可以通过构象的变化堵住通道。在干旱的时候,通道因为两个保守的丝氨酸残基发生脱磷酸化而关闭。而在水涝的时候,氧气的缺乏导致细胞质基质的 pH 下降,也促使通道上一个关键的组氨酸残基因质子化而关闭。两种情况都是环境因素的变化而诱发蛋白质的构象发生变化,而构象变化促使肽环移至通道口并堵住通道。

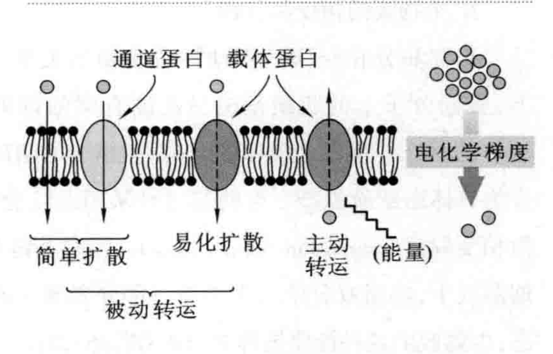


图 16-25 生物膜的转运系统

主动转运是一种直接依赖于能量并且逆浓度梯度进行的转运方式。能量的来源可能直接偶联于 ATP 的水解,或者偶联于某一离子顺浓度梯度的转移。其中,直接偶联于 ATP 水解的主动转运方式被称为初级主动转运(primary active transport),而偶联于某一离子顺浓度梯度转移的主动转运方式被称为次级主动转运(secondary active transport)。主动转运可导致一种物质在膜一侧的积累,因此对细胞的生命活动具有特别重要的意义。

Quiz7 如果人为地让一个细胞内的 Na^+ 浓度高于细胞外,那么 Na^+ 通过 Na^+/K^+ 泵离开细胞还需要消耗 ATP 吗?为什么?

生物大分子的跨膜转运采取与小分子物质完全不同的策略,进入细胞通过内吞(endocytosis)或胞吞,离开细胞则通过胞吐(exocytosis)。其中内吞的基本过程是:局部细胞膜内陷→包被附近的大分子→形成小囊泡→囊泡脱离细胞膜,进入胞内→囊泡与细胞内膜融合(通常是溶酶体),将内容物转运到目的地。完成内吞一般需要消耗能量,还需要细胞质基质中的钙离子以及细胞内运动系统的协助。

e16-3 主动转运
e16-4 内吞与胞吐

科学故事——水孔蛋白的发现

生命系统约 70% 由水组成,所有生物都需要水进出细胞。科学家很早就知道水分子除了能够以简单扩散的手段通过细胞膜以外,还应该存在其他的机制,因为许多细胞对水的通透性要比简单扩散快得多,而且如果水仅仅通过被动转运机制进出细胞,那么渗透压很容易导致细胞破裂或者细胞脱水(取决于细胞膜内外的盐浓度)。但这种机制究竟是什么却一直悬而未决,直到约翰·霍普金斯大学的 Peter Agre 在细胞膜上发现了水孔蛋白。

水孔蛋白发现的故事再次验证了"运气只会眷顾那些有准备的人"这句至理名言。

在 20 世纪 80 年代中期,Agre 及其同事在红细胞膜上寻找作为 Rh 因子一部分的蛋白质,偶然得到一种含量丰富、非常小的蛋白质"不速之客"。他们很快分离到这种蛋白质,并将它命名为 CHIP28,而且发现它也存在于肾细胞的质膜上。很快,他们用了不到一年的时间测定出它的氨基酸序列并克隆到其 cDNA。考虑到这种蛋白质也存在于与水代谢密切相关的肾细胞的质膜上,于是他们就想:也许它就是组成水通道的蛋白质!

为了证明这种推测,他们构建了 CHIP28 表达质粒,随后,把构建好的质粒显微注射到非洲爪蟾(*Xenopus laevis*)的卵里进行表达。当他们将含有 CHIP26 表达质粒的卵放到低渗介质中以后,发现细胞迅速发生膨胀,而没有 CHIP26 表达质粒的卵形状没有变化(图 16-26)。此外,他们还将 CHIP26 重组到脂质体上,结果也发现这种人造细胞也能够从低渗溶液中吸水膨胀。

当他们将 Hg^{2+} 与 CHIP26 混在一起的时候,发现 Hg^{2+} 能使 CHIP26 失活,这种结果解释了 Hg^{2+} 能够阻止水的跨膜转运这个很早就为人所知的现象。综合以上的实验结果,可以毫不怀疑地认为,他们发现

的 CHIP26 就是水孔蛋白。

2000 年，Agre 和其他几位科学家得到了红细胞膜水孔蛋白（现在被称为水孔蛋白 -1，AQPI）的三维结构（图 16-27）。水通道的三维结构清楚地表明，其肽链两个部分形成的半孔（hemipore）组装成了一个允许水分子通过的通道。

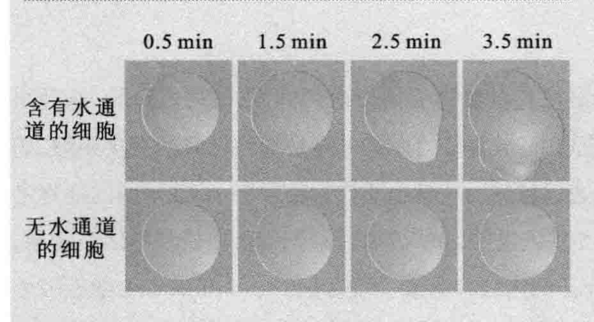


图 16-26 表达与不表达 CHIP26 的非洲爪蟾卵对水通透性的差别

水通道的发现迎来了对这种存在于各种生物的蛋白质进行生化、生理学和遗传学等研究的黄金时代，对于它的研究可以更好地揭示生命现象，使科学家能够研发出一些治疗因水通道异常而导致的疾病。Agre 因此荣获 2003 年的诺贝尔化学奖。

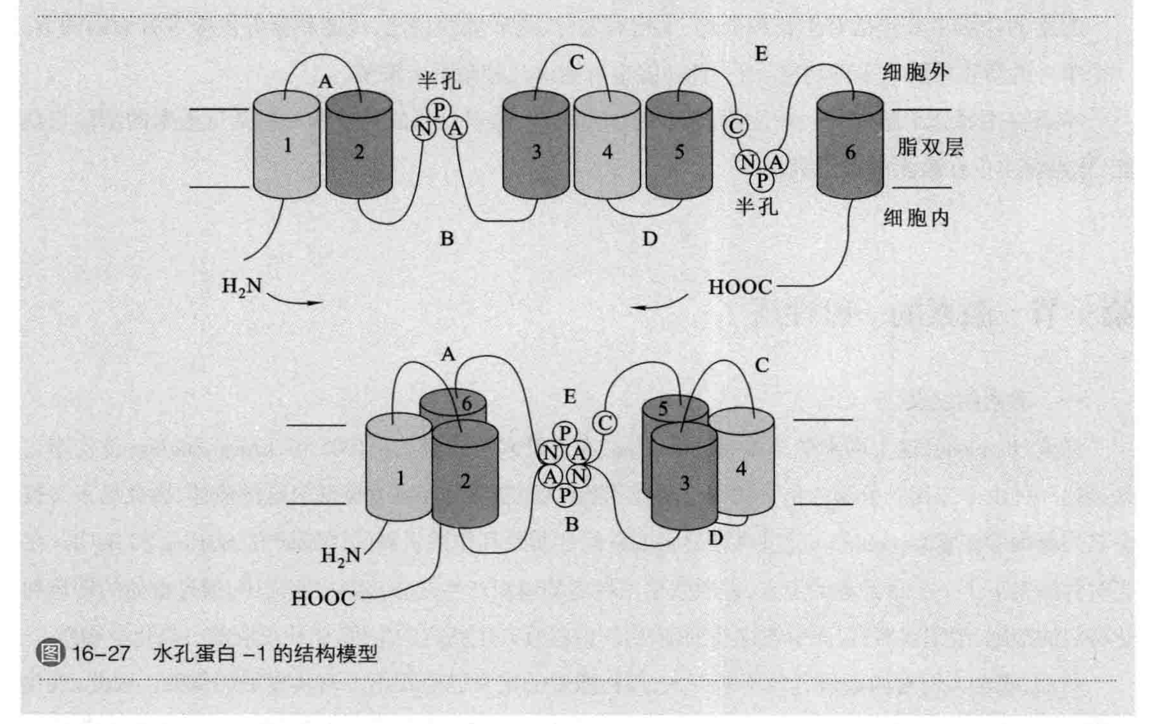


图 16-27 水孔蛋白 -1 的结构模型

推荐网址：

e16-5 本章小结填填看

1. http://en.wikipedia.org/wiki/Lipid（维基百科有关脂质的内容）
2. http://themedicalbiochemistrypage.org/lipids.html（完全免费的医学生物化学课程网站有关脂质的内容）
3. http://www.sci.sdsu.edu/classes/biology/bio202/TFrey/MembraneStructure.html（美国圣地亚哥州立大学一个生物学网站有关生物膜结构的内容）
4. http://www.bms.ed.ac.uk/research/others/smaciver/Cyto-Topics/lipid_rafts_and_the_cytoskeleton.htm（英国爱丁堡大学生物医学院一个网站有关最新的脂筏结构的内容）

参考文献：

1. Linton K J, Structure and function of ABC transporters. *Physiology*, 22(2): 122–130, 2007.
2. Zajchowski L D, Robbins S M. Lipid rafts. *Eur J Biochem*, 269, 737–752, 2002.
3. Cary L A, Cooper J A. Molecular switches in lipid rafts. *Nature*, 404: 945–947, 2000.
4. Walz, et al. The 6Å three-dimensional structure of AQP1. *Nature*, 387: 624–627, 1997.

第十七章　激素及其受体介导的信号转导

多细胞生物能够生存并有效地行使功能，组成它们的细胞必须以协调的方式进行活动。这种协调需要在不同的细胞之间进行通讯。细胞通讯最简单、最直接的方式可能是细胞之间的直接接触，如间隙连接(gap junction)。而比较复杂的方式则是通过特殊的化学物质作为信号分子，在不同的细胞之间进行信号转导(signal transduction)。据估计，人类基因组约 20% 的基因编码的蛋白质参与信号转导。此数据突显了信号转导的重要性。能充当信号分子的物质有激素、生长因子(growth factor)、细胞因子(cytokine)和神经递质(neurotransmitter)等。无论是哪一种信号分子，都要先与特定细胞上的专一性受体(receptor)结合，然后才能发挥作用。

激素的功能主要包括对生长和分化、生殖和发育、内环境的稳定、代谢和营养供应等方面的调节。一个单一的激素可能有多种功能，而一种功能也可能由几种激素来调节。

本章将主要介绍激素的一般性质、激素的作用机理、激素分泌的调节和一些常见激素的结构与功能，重点将集中在激素的作用机制。

第一节　激素的一般性质

一、激素的定义

激素(hormone)这个词来源于希腊语 hormao，本意是兴奋或激动。1905 年，Ernest Starling 首先用它来描述一种由十二指肠分泌的化学物质，该物质能够刺激胰腺分泌富含碳酸氢盐的液体，因此后来又被命名为肠促胰液素(secretin)。经过对肠促胰液素的性质及其作用机制的详细研究，Starling 和 Bayliss 在当时为激素下了一个最经典的定义，即激素是由特定的组织产生并分泌到血流之中，通过血液的转运到达特定的细胞、组织或器官，而引发这些细胞、组织或器官产生特定的生理、生化反应的一类化学物质。

然而，随着人们对激素研究的不断深入，最初激素的定义已经跟不上时代发展的需要。因此，现代的激素定义被赋予了更广泛的含义。现一般认为，激素是一类非营养的、微量就能起作用的在细胞间传递信号的化学物质。就动物而言，分泌激素的细胞称为内分泌细胞，受激素作用的细胞称为靶细胞。在以上定义的基础上，还可以根据作用的距离，将动物激素进一步分为内分泌(endocrine)激素、神经内分泌(neuroendocrine)激素、旁分泌(paracrine)激素、自分泌(autocrine)激素和内部分泌(intracrine)激素，这 5 类激素的差别参见图 17–1。其中内分泌激素和神经内分泌激素的作用距离最远，且绝大多数激素属于这一类；旁分泌激素只作用于临近的细胞，作用时间短，例如前列腺素、阿片肽(opioid peptide)以及一些多肽生长因子；自分泌激素则作用于原来分泌它的细胞，如刺激 T 细胞分裂的白介素 –2(interleukin–2)、某些生长因子以及少数细胞癌基因的产物；内部分泌激素在细胞内合成以后不需要分泌到胞外，而是留在原来的细胞内起作用。按照更为广泛的激素定义，生长因子、细胞因子和神经递质等许多重要的细胞间信号分子都可归为激素。

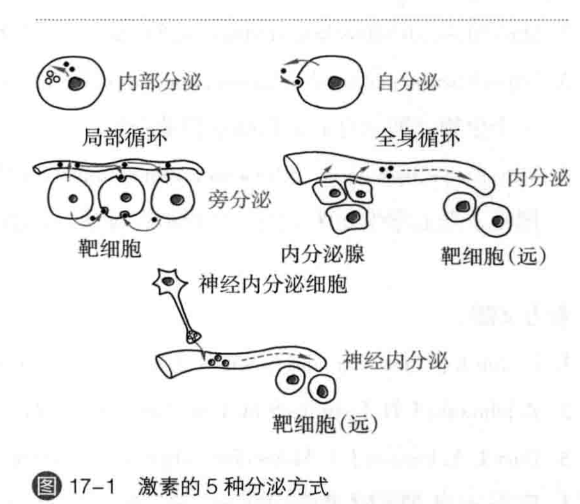


图 17–1　激素的 5 种分泌方式

二、激素的化学本质、分类和生物合成

激素的种类繁多，彼此之间的结构差别可能很大，但根据其来源和化学本质，可将它们分为肽类或蛋白质激素、固醇类激素、氨基酸衍生物激素、脂肪酸衍生物激素等几类。

肽类或蛋白质激素的种类最多，小到仅由 3 个氨基酸组成，大到由几百个氨基酸组成。

固醇类激素均衍生于胆固醇，包括维生素 D 以及由肾上腺皮质和生殖腺分泌的皮质激素和性激素。所有固醇类激素都具有相同的核心环结构，虽然它们的表面二维结构十分相似，但是各自的侧链和在空间的取向却不一样，这就决定了各种固醇类激素作用的特异性。

氨基酸衍生物激素衍生于特定的氨基酸，如 Tyr 或 Trp。衍生于 Tyr 的激素肾上腺素(epinephrine)、去甲肾上腺素（norepinephrine）和甲状腺素（thyroid hormone），衍生于 Trp 的激素有血清素（serotonin）和褪黑素（melatonin）。

脂肪酸衍生物激素衍生于脂肪酸，如衍生于花生四烯酸的激素包括前列腺素（prostacyclin）、凝血恶烷（thromboxane）和白三烯（leukotriene）。

按照溶解性质，可以将激素简单地分为水溶性激素和脂溶性激素。这样分类的好处是可以更容易理解激素的作用机制。这两类激素的主要差别有 3 点：

(1) 脂溶性激素很容易通过脂溶性的生物膜，因此难以贮存在胞内，一般在需要的时候才被合成，但甲状腺素是一个例外。而水溶性激素可以被包被在具有膜结构的囊胞内，在体内贮存方便，当需要的时候可立即分泌出去。

Quiz1 你认为甲状腺素例外的原因是什么？

(2) 脂溶性激素很难溶于水，在动物体内需要与血清中特殊的转运蛋白结合后才能转运。这种结合反过来又保护了激素，提高了激素的稳定性。除了一些小肽，绝大多数水溶性激素的转运并不需要与血清蛋白结合，这就使得它们很容易被代谢掉。

(3) 脂溶性激素的疏水性质允许它们自由通过细胞膜，与胞内的受体结合，产生细胞内效应。但水溶性激素不能通过质膜，因此它们必须与细胞膜上的受体结合才能发挥作用。

绝大多数蛋白质和肽类激素由单一的基因编码，少数由两个或两个以上的基因编码。与其他分泌蛋白一样，多肽激素以前体的形式被合成。肽类激素最初的形式通常是前激素原（pre-prohormone）。前激素原除了在 N 端含有指导其进入糙面内质网的“前序列”（信号序列）以外，还含有多余的“原序列”，这“一前一原”的序列最后都要被切除。前激素原在进入内质网以后被切除“前序列”，变成激素原（prohormone），激素原在高尔基体再被切掉“原序列”而转变成有功能的激素，并被包装和贮存在衍生于高尔基体的分泌小泡内，随时准备分泌到胞外。

固醇类激素的合成主要发生在光面内质网，但有少数反应发生在线粒体。一种细胞要合成某种固醇类激素，必须存在相关的酶。

氨基酸衍生物激素主要是通过对氨基酸的侧链修饰而成，而脂肪酸衍生物激素主要由花生四烯酸在特定的酶催化下形成。

三、激素的定量

精确、快速地测定激素的浓度对于激素的研究十分重要，但由于机体内各种激素的浓度都很低，要确定一种激素的浓度并非易事。早期人们使用化学分析法（chemical assay）和生物活性测定法（bioactivity assay）来测定激素的浓度，其中前一种方法是根据激素的化学性质来进行化学定量，后一种方法是根据激素的生物活性（比如胰岛素能降血糖）来定量。显然，这两种方法都很难精确、有效和特异性地测定出某一种激素的浓度。就化学分析法而言，许多激素的反应性质一样，无特异性可言。而生物活性测定法对于很多激素来说，根本就行不通。例如，下丘脑激素的生物活性是促进或抑制脑垂体激素的分泌，这样的生物活性是很难定量的。再说，体内有很多激素具有相同或相似的生物活性(例如肾上腺素、胰高血糖素和糖皮质激素都可以升血糖)。由于上述两种方法的严重缺陷，早期人们对激

素的研究一直没有显著突破。直到 20 世纪 50 年代美国科学家 Rosalyn Sussman Yalow 创立了放射免疫测定法(radioimmunoassay, RIA)以后,激素的研究才进入了突飞猛进的黄金时代。基于 Yalow 对激素研究的杰出贡献,她荣获了 1977 年的诺贝尔生理学或医学奖。

Quiz2 哪一种激素无法用 RIA 测定浓度?

RIA 利用了抗原(antigen, Ag)与抗体(antibody, Ab)之间能够以高度特异性的方式可逆结合的性质。如图 17-2 所示,假如以一种激素(如 ACTH)为抗原,其中一部分用放射性同位素标记(*Ag),另一部分不进行标记(Ag),当将它们混合在一起与抗体保温的时候,两者与抗体的亲和性完全相同,但有竞争性。显然,同位素标记的激素越多,形成的具有放射性的激素 – 抗体复合物(Ab-*Ag)的量就越多。假定最初与抗体一起保温的激素都被同位素标记的话,那么与抗体结合的将完全是被同位素标记的激素。随后若再依次加入一定量的非同位素标记的同种激素,加入的非标记的激素就会取代一部分与抗体结合的同位素标记的激素,且加入的非标记的激素量越多,被取代的同位素标记的激素就越多。根据这一点,就可以以加入的[非同位素标记的激素]为横坐标(图中为 ACTH),[与抗体结合的同位素标记的激素]/[游离的同位素标记的激素]为纵坐标,然后做标准曲线。未知量的激素在与同一种反应体系的抗体结合以后,可以根据标准曲线求得其浓度。

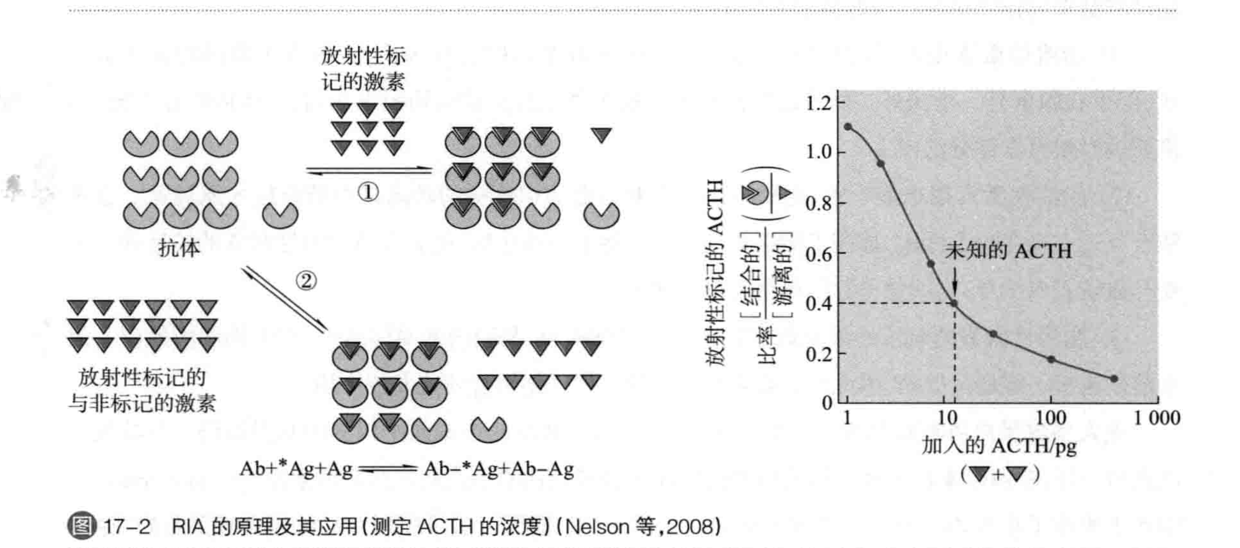


图 17-2 RIA 的原理及其应用(测定 ACTH 的浓度)(Nelson 等, 2008)

第二节 激素作用的一般特征

激素的作用具有特异性、高效性,往往需要"第二信使",可产生快、慢反应,还具有脱敏性和时效性等性质,其中以特异性最为重要。

一、特异性

激素作用的特异性是指一种激素只能作用一种或一类细胞的现象。当一种激素由特定的细胞分泌后,可随着体液循环与机体内几乎所有的细胞接触,但为什么只有特定的细胞才有反应呢?这是因为激素的作用离不开受体,一种激素只有在与其专一性受体分子结合以后才能发挥作用,而只有特定的细胞才存在特定激素的受体。那什么又是受体呢?

(一) 受体的基本概念

受体是存在于细胞之中的一种特殊成分,它能够识别并结合源自细胞外的各种信号配体,形成可逆的二元复合物,由此引发出特定的生物学效应。从化学本质上来看,受体主要是蛋白质,尤其是糖蛋白,一些糖脂也可充当受体。当一个配体与其受体结合以后,会诱导受体的构象发生变化,并在此基础上诱发靶细胞产生特定的生理、生化反应。

（二）受体的基本性质

体内和体外的一系列研究表明，受体至少具有以下一些基本性质：

（1）与配体结合的高度专一性　受体和配体的结合与酶和底物的结合很相似，都表现出高度的专一性。例如，胰岛素只会与自身的受体结合，而不会与胰高血糖素的受体结合。当然，这种专一性是相对的，如肾上腺素就有5种不同的受体（α_1、α_2、β_1、β_2 和 β_3）。同一种激素所具有的在结构和功能上有区别的几种受体称为同工受体（isoreceptor）。同一种激素与不同的受体结合可能产生不同的效应，有时甚至是相反的效应。

（2）与配体结合的可逆性　受体和配体的结合方式是以非共价键（氢键、疏水键和离子键等）相连，因此是完全可逆的。这一性质有利于激素作用的终止。

（3）与配体结合的高亲和性　受体与配体之间的亲和力很强，亲和常数（K_a）高达 10^9~10^{10} mol/L，因此，虽然激素在体内的浓度很低，但仍然能够识别并结合相应的受体。

（4）与配体的结合可产生强大的生物学效应　激素与受体结合后，形成二元复合物，由此启动信号转导，最终导致靶细胞产生特殊的生物学效应。10^{-9} mol/L 的激素可导致靶细胞内代谢物浓度发生 10^6 倍的变化，可见激素诱发的生物学效应是巨大的。

框 17-1　生化与健康——瘦肉精、喘息定与萘心安

瘦肉精、喘息定与萘心安这3种物质在结构上都与肾上腺素类似（图17-3），但进入体内以后，有的作为肾上腺素的激动剂起作用，有的是作为肾上腺素的颉颃剂起作用。激动剂（agonist）在结构和功能上均能够模拟激素，既能与受体结合，又能诱发强大的生物学效应，而颉颃剂（antagonist）仅仅能在结构上模拟激素，相当于酶的竞争性抑制剂，与受体结合以后并不能诱发任何生物学效应的产生。

瘦肉精　异丙肾上腺素　肾上腺素　萘氧丙醇胺

图 17-3　肾上腺素及其激动剂和颉颃剂的结构比较

喘息定即异丙肾上腺素（isoproterenol）就是一种激动剂，它能够与肾上腺素的 β_2 受体结合，松弛支气管，因此可以作为治疗哮喘的药物。瘦肉精即克伦特罗（Clenbuterol），也是一种激动剂，能够与脂肪细胞膜上的肾上腺素 β_2 受体结合，促进脂肪细胞内脂肪的分解，有利于瘦肉的生长。它曾经作为药物用于治疗支气管哮喘，后由于副作用太大而遭禁用。瘦肉精被过量使用后容易残留在动物内脏特别是肝上，在上海曾经引发了几百人的中毒事件。由于西方人一般不食用动物内脏，因而，在美国、加拿大、新西兰等国，瘦肉精这类物质的使用是合法的。萘心安即萘氧丙醇胺（propranolol，普萘洛尔）却是一种颉颃剂，它与 β_2 受体结合以后，不产生任何效应，但能抑制肾上腺素的作用，因此可以作为治疗心律失常的药物。

（5）与配体结合的饱和性　一个细胞上特定受体的数目是有限的，因而激素与受体的结合具有饱和性动力学特征。当激素浓度达到一定的水平时，所有的受体都被激素占据，也就达到了饱和状态。当然，细胞上一种受体的数目并不是恒定的，在特殊的生理或病理条件下，受体数目会发生变化。调节受体数目的主要原因是激素本身，一般情况下，激素浓度的提高或激素长时间与靶细胞接触都可引起受体数目的下调（down-regulation）。

由此可见，受体在与相应的配体相互作用的时候，具有类似于米氏酶与底物结合时表现的一些特性。为了更好地显示两者结合的动力学性质，可使用 Scatchard 作图法来研究激素与受体的相互作用。

e17-1 Scatchard 作图

通过这种作图的方法，不仅可以测定出一个靶细胞上特定激素受体的数目，还可以得到激素和受体结合的亲和常数 K_a。

（三）受体的分类和结构

激素受体的种类和结构都比较复杂，根据受体在细胞中的定位，可将它们分成细胞膜受体和细胞内受体，而后者又可分为细胞质受体和细胞核受体。水溶性激素由于不能通过生物膜，其受体都位于细胞膜上；而脂溶性激素很容易通过细胞膜，因此受体大多数位于细胞内，但若是位于细胞膜上，也不会影响它与激素的结合，故有少数脂溶性激素的受体就位于细胞膜上。例如，孕酮在非洲爪蟾内的受体就位于其卵母细胞的细胞膜上。再如，油菜素内酯（brassinosteroid，BR）是一类重要的脂溶性植物激素，其受体也在细胞膜上。正是因为两类激素受体在细胞中的位置不同，作用机制才会有较大的差异。

1. 细胞质膜受体

细胞质膜受体又名细胞表面受体或膜受体。细胞质膜是一种脂双层结构，因此所有的膜受体都具有一个或几个疏水区。膜受体的疏水区“深陷”在细胞膜内使其很难分离，故分离纯化膜受体需要使用两性去垢剂（如脱氧胆酸盐）来增溶。有的膜受体并不贯穿整个细胞膜，而有的却穿过质膜内外，甚至不止一次。受体的跨膜区一般富含疏水氨基酸，常形成 α 螺旋。但不管怎样，所有膜受体上的配体结合部位都应该面向胞外，因为只有这样，受体才会被配体识别并结合。

膜受体主要分为 4 类。这 4 类膜受体在结构和功能上有明显的差异：

（1）与 G 蛋白偶联的受体（G protein-coupled receptor，GPCR） 此类受体一般都具有 7 次跨膜的疏水 α 螺旋（图 17-4），如肾上腺素的 β 受体，因此又名 7TM 受体。它的功能与 G 蛋白紧密偶联（G 蛋白的结构与功能见后），其中与 G 蛋白作用的结构域位于细胞质基质一侧第 5 个和第 6 个 α 螺旋之间的环上。这类受体构成了一个庞大的蛋白质超家族。据估计，人类基因组有 800 多个基因在编码这个家族中的各种成员，同时多种疾病与它们有关系，而现代临床用的药物约有 30% 以它们作为靶标。

（2）离子通道受体（ion-channel receptor） 离子通道是位于细胞膜上的水溶性通道，其功能是允许或阻止离子进出细胞。离子通道的默认状态通常是闭合，受到特定信号的刺激以后就会开放（图 17-5）。

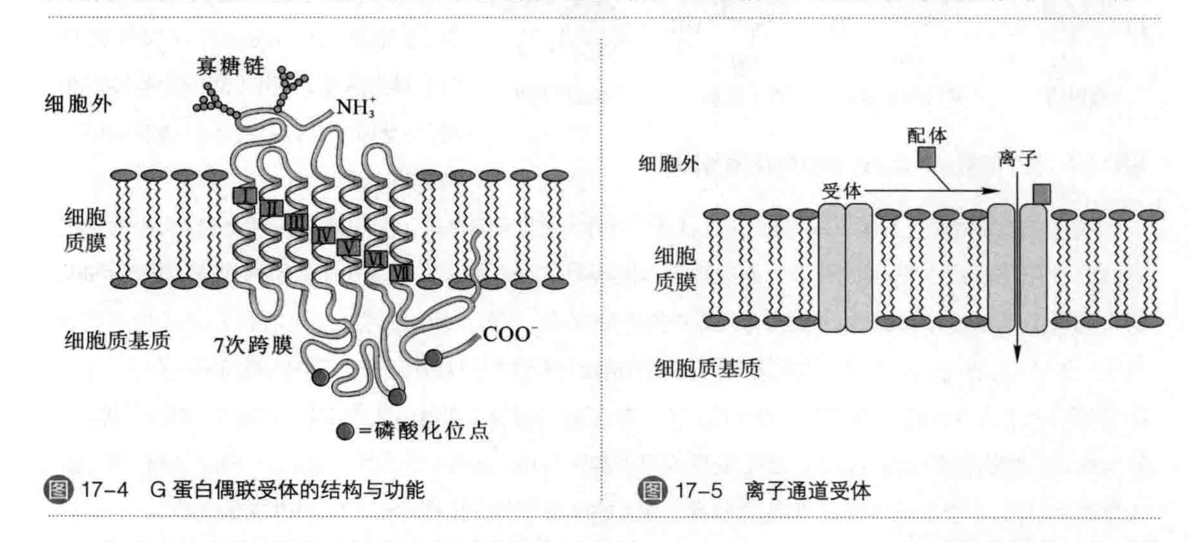


图 17-4 G 蛋白偶联受体的结构与功能

图 17-5 离子通道受体

有三种类型的离子通道：①电位门控的离子通道（voltage-gated channel），打开此类通道的信号是与动作电位（action potential）相关联的膜电位的变化；②第二信使门控的离子通道（second-messenger-gated channel），打开此类通道的是特殊的第二信使（见后），如 cAMP 和 cGMP；③配体门控的离子通道（ligand-gated channel），此类通道本身就是受体，打开它的信号是配体的结合。例如，乙酰胆碱的烟碱型受体（nicotinic receptor）就是 Na^+ 通道（图 17-6），其组成为 $\alpha_2\beta\gamma\delta$。每一个亚基都要跨膜 4 次，跨膜部分为 4 段 α 螺旋，其中一段 α 螺旋含较多的亲水氨基酸，正是由于这个亲水区的存在，使得 5 个亚基在膜上能够组装成一个亲水性的通道，而乙酰胆碱结合在 α 亚基上。再如，γ- 氨基丁酸（γ-aminobutyric acid，GABA）的 A 型受体为 Cl^- 通道，血清素的受体为 Na^+ 和 K^+ 通道。

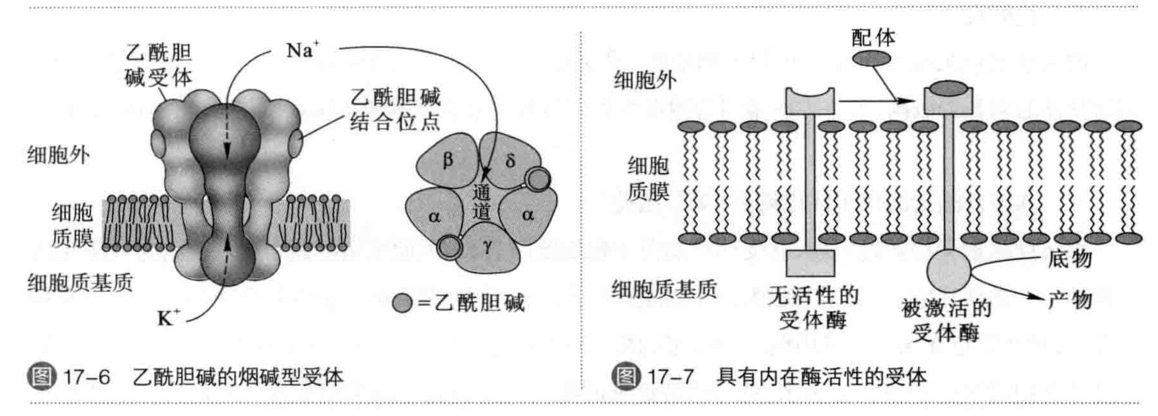


图 17-6　乙酰胆碱的烟碱型受体

图 17-7　具有内在酶活性的受体

(3) 酶受体　此类受体有潜在的酶活性(图 17-7),如心房利钠肽和胰岛素的受体分别具有鸟苷酸环化酶和酪氨酸蛋白质激酶的活性,转化生长因子 β(transforming growth factor β,TGF-β)的受体具有丝氨酸蛋白质激酶的活性,白细胞共同抗原 CD45 受体(leukocyte common antigen CD45 receptor)具有蛋白质酪氨酸磷酸酶的活性。当这些受体与配体结合后,酶活性即被激活。

(4) 无酶活性但直接与细胞质内酪氨酸蛋白质激酶相联系的受体　此类受体与受体酪氨酸激酶相似,但本身没有任何酶活性(图 17-8)。干扰素(interferon)和人生长激素的受体均属于这一类。

以上 4 类受体并不能覆盖所有激素的膜受体。事实上,某些激素的膜受体很难按照这 4 类来对号入座。如与细胞凋亡有关的肿瘤坏死因子(tumor necrosis factor,TNF),其受体就不属于这 4 类,但由于这种受体在细胞质一侧有一种特殊的结构域——死亡结构域(death domain),故名为死亡结构域受体。

2. 细胞内受体

激素的细胞内受体又可以分为细胞质基质受体和细胞核受体,例如,醛固酮的受体位于细胞质基质,而甲状腺素、1,25-$(OH)_2VD_3$ 和视黄酸等其他一些脂溶性激素的受体则位于细胞核。有一些细胞内受体的配体到目前还不能确定,这些受体经常被称为孤儿受体(orphan receptor)。

细胞内受体至少含有 3 个活性部位:一个负责与激素结合;另一个负责与 DNA 上特殊的被称为激素应答元件(hormone response element,HRE)的碱基序列结合,此活性部位有一种特殊的结构模体——锌指结构(zinc finger)(图 17-9)(参看第四十三章“真核生物的基因表达调控”);第三个活性部位是让受体之间形成同源二聚体。

如果是细胞质基质受体,则还有第四个活性部位。在没有受体的时候,该部位与热激蛋白 90(HSP90)以及热激蛋白 70(HSP70)结合。HSP90 又与免疫亲和蛋白(immunophilin,IP)结合。当受体与激素结合后,受体构象发生变化,从而导致 HSP90、HSP70 和 IP 的释放。

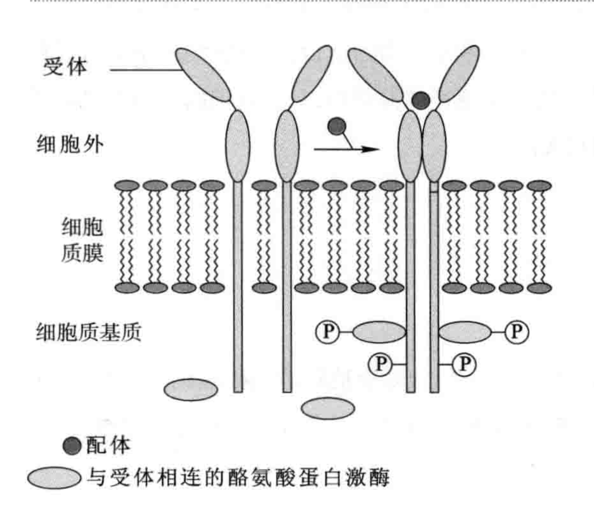


图 17-8　无酶活性但直接与细胞质内的酪氨酸蛋白质激酶相联系的受体

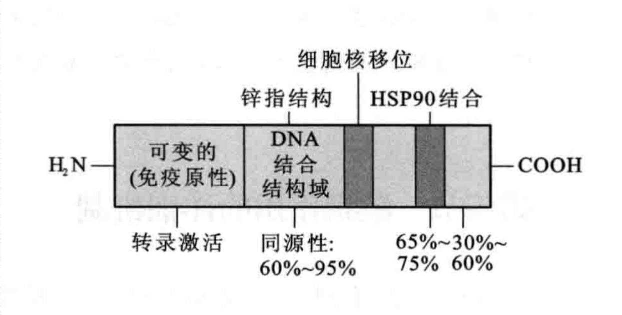


图 17-9　固醇类激素细胞质基质受体结构模型

二、高效性

激素在很低的浓度下就能作用于靶细胞,诱发靶细胞产生强烈的生物学效应。一方面是因为激素与受体的亲和力极高,另一方面是因为激素在作用过程中存在级联放大(cascade amplification)机制。

三、水溶性激素的作用往往需要“第二信使”

水溶性激素无法通过质膜,其受体只能位于靶细胞的表面,因此水溶性激素无法直接进入细胞内发挥作用。然而当它们与靶细胞膜上相应的受体结合后,形成的激素-受体复合物却可以通过某种手段,去激活定位在细胞质膜内侧一种特定的酶,从而导致一些小分子物质的合成。这些小分子物质一旦从酶的活性中心被释放出来,就可代替原来的激素行使功能。如果把激素看成是第一信使(first messenger),随后合成的小分子物质就可以看成是第二信使(second messenger),而催化第二信使合成的酶可称作效应器(effector)。到目前为止,已发现的第二信使有:cAMP、cGMP、1,4,5-三磷酸肌醇(inositol 1,4,5-triphosphate,IP_3)、Ca^{2+}、二酰甘油(DG)、神经酰胺、花生四烯酸、NO、CO 和 H_2S 等。

四、可能产生“快反应”或“慢反应”

有的激素作用靶细胞后,靶细胞在很短的时间内就能产生特殊的生物学效应,这种现象称为“快反应”(acute effect);相反,有的激素作用靶细胞后,靶细胞需要较长的时间才能产生特殊的生物学效应,这种现象称为“慢反应”(chronic effect)。激素“快反应”的特征是“来也匆匆,去也匆匆”,在几秒钟或几分钟内产生效应,反应强度与激素浓度成正相关,反应持续的时间通常不长。而激素“慢反应”的特征是“来得慢,去得也慢”,在几小时甚至几天后才能产生明显的效应,反应强度与激素浓度在短时间内的波动无关,反应持续的时间较长。

水溶性激素一般产生“快反应”,而脂溶性激素通常产生“慢反应”,还有一些激素既可产生“快反应”,又可产生“慢反应”,如胰高血糖素和胰岛素。一种激素究竟是产生“快反应”还是产生“慢反应”,与它的作用机制直接有关。

五、脱敏性和时效性

一种激素在与其靶细胞长时间接触后,靶细胞倾向于降低其反应性,这种现象称为激素的脱敏作用(desensitization)。产生脱敏的原因可能是受体数目的下调或受体经历了共价修饰(如磷酸化)。

激素的作用具有一定的时效性。一种激素的作用,从分泌到与受体结合,再到产生最终的生物学效应不可能持续很久,否则,靶细胞会出现功能紊乱甚至死亡。激素作用的时效性产生的原因是靶细胞内存在各种信号终止的机制,这些机制可在不同的环节发挥作用,保证了靶细胞不会持续地受到一种激素作用,同时也为靶细胞接受下一轮激素的作用做好了准备。实际上,每当细胞的一条信号通路启动以后,该通路上某一个被激活的成分通常又作为同一条途径的抑制剂,只不过抑制剂的作用略滞后于激活剂,这样保证了在信号发送之前不至于就被关闭。

第三节 激素作用的详细机制

激素的作用过程一般包括以下几步:①激素的合成和分泌;②激素被转运到靶细胞;③激素与靶细胞膜或靶细胞内的特异性受体结合,致使受体被激活;④靶细胞内的一条或几条信号通路被启动;⑤靶细胞产生特定的生物学效应;⑥信号的终止。

下面主要介绍信号途径的启动和具体的转导过程。

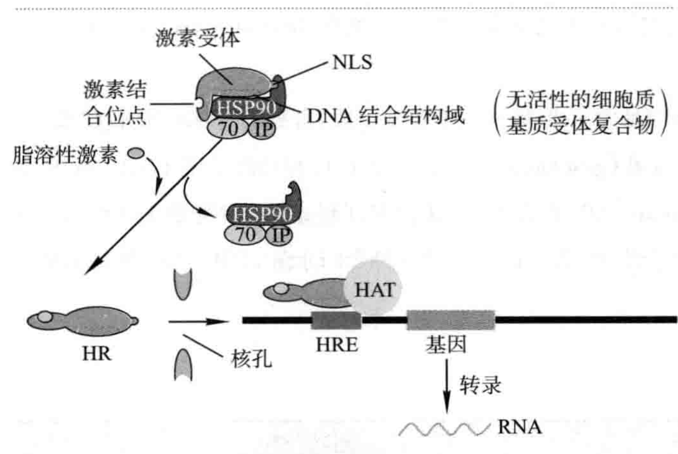


图 17-10 受体位于细胞质的脂溶性激素的作用机制

一、脂溶性激素的作用机制

正如前述，脂溶性激素的受体通常位于细胞内，因此脂溶性激素一般必须进入细胞内才能起作用。而那些受体位于细胞质基质的脂溶性激素就很容易自由扩散到细胞质基质，并与其中的受体结合。这种结合改变了受体的构象，使原本与受体结合的HSP90、HSP70和IP被释放（图17-10）。一旦HSP90、HSP70和IP得到释放，原先在受体上被隐藏的细胞核定位信号（nuclear localization signal，NLS）以及负责与HRE结合的结构域就暴露出来。在NLS的指导下，激素与受体的二元复合物（HR）从核孔进入细胞核（第四十章“蛋白质的翻译后加工及其定向和分拣”），在形成二聚体——$(HR)_2$以后，通过受体上的锌指结构与DNA分子上高度特异性的HRE结合，再将组蛋白乙酰转移酶（histone acetyltransferase，HAT）招募进来，催化HRE周围染色质上的组蛋白发生乙酰化修饰，促使局部的染色质从不利于基因表达的紧密构象变成有利的松散构象。不同脂溶性激素的HRE序列存在差异（参见表17-1）。

表 17-1 常见的脂溶性激素 HRE 一致序列

激素	HRE 一致序列	激素	HRE 一致序列
雌二醇	5′GGTCANNNTGACC3′	孕酮	5′GNANANGNTGTYC3′
糖皮质激素（刺激性）	5′GNACANNNTGTYCT3′	甲状腺素（T3）	5′AGGTAAGATCAGGGACG3′
糖皮质激素（抑制性）	5′CAGGAAGGTCACGTC CAAGGGCTC3′		

对于受体位于细胞核的脂溶性激素来说，需要先后跨过靶细胞的质膜和核膜，才能到达核内与相应的受体结合。在没有激素的时候，这些“闲置”的细胞核受体有的也与HSP90结合（如雌激素的受体），但以单体形式存在。然而，一旦有激素与之结合，HSP90即解离下来。此时受体形成二聚体，并与HRE结合，再通过招募HAT激活特定的基因表达。

还有一些细胞核受体在没有激素的时候，就已结合在相应的HRE上（如甲状腺素的受体），同时还与组蛋白去乙酰酶（histone deacetylase，HDAC）结合。与HAT正好相反，HDAC保持局部的染色质处在浓缩的构象，而阻止附近的基因的表达。一旦有激素结合，HDAC立刻被释放。以此同时，HAT则被招募进来，并催化组蛋白的乙酰化修饰，促使局部的染色质从紧密构象变成松散开放的构象，这有利于催化基因转录的RNA聚合酶和转录因子结合到启动子上，从而激化HRE下游基因的表达（参看第三十六章“DNA转录”）。

Quiz3 如果将雌二醇的HRE倒置，它还能控制原来的基因表达吗？为什么？

二、水溶性激素的作用机制

水溶性激素的种类最为繁多，作用机制也最复杂，为了叙述上的方便，可以按照受体的性质或第二信使的特征将它们划分为以下几类：

（一）与G蛋白偶联的受体系统

该系统需要一种特殊的蛋白质——G蛋白，G蛋白的发现以及结构与功能的鉴定主要归功于美国的Alfred Gilleman和Martin Rodbell。这两位科学家因此荣获了1994年的诺贝尔生理学或医学奖。

G蛋白就是鸟苷酸结合蛋白，它在系统中的功能是作为一种中间接受体（acceptor），在受体和效应

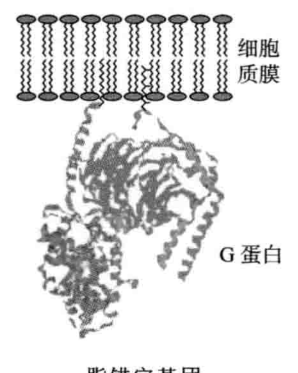
器之间转导信号。已发现两类不同的G蛋白,一类是异源三聚体G蛋白(heterotrimeric G protein),另一类是小G蛋白。

Quiz4 是不是第二信使cAMP只能通过激活PKA才能起作用?

(1) 异源三聚体G蛋白 这一类G蛋白是由α、β和γ三个亚基组成,常见的有:Gs蛋白、Gi蛋白、Gq蛋白、Gt蛋白、Golf蛋白、Go蛋白和味觉素(gustducin),它们的主要特征和功能见表17-2。其中Gs蛋白能够刺激腺苷酸环化酶(adenylate cyclase,AC)的活性,Gi蛋白则抑制腺苷酸环化酶的活性,Gq能够激活磷脂酶C(phospholipase C,PLC)的活性,Gt蛋白能够激活一种专门水解cGMP的磷酸二脂酶的活性,Golf蛋白也能够刺激AC的活性。

表17-2 常见的异源三聚体G蛋白的比较

G蛋白	效应器	第二信使	受体实例
Gs	AC	cAMP(↑)	肾上腺素β受体,加压素受体,血清素受体,胰高血糖素受体
Gi	AC	cAMP(↓)	肾上腺素的α_2受体
Go	K^+通道(βγ激活)	膜电位变化	乙酰胆碱毒蕈碱型受体
Golf	AC	cAMP(↑)	嗅觉受体细胞膜上的有味物质受体
Gq	PLC	DG,IP_3(↑)	肾上腺素的α_1受体,GnRH受体,乙酰胆碱在内皮细胞的受体
Gt	G-PDE	cGMP(↓)	视紫红质



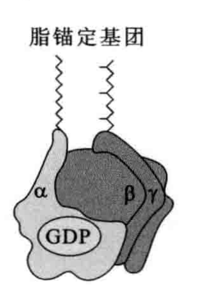


图17-11 G蛋白的结构组成以及它与细胞质膜内侧的结合

异源三聚体G蛋白的结构参见图17-11。以Gs蛋白为例,它属于第一类脂锚定蛋白(参看第十六章“脂质与生物膜”),其α亚基在N端因豆蔻酰化而被封闭,Gs蛋白正好通过这个脂溶性的豆蔻酰基锚定在细胞膜内侧。体外的研究表明,α亚基能结合激素的受体、腺苷酸环化酶、氟化物(一种外源的激活物)和GTP。此外,α亚基还有缓慢的GTP酶活性,且能受霍乱毒素的修饰。β亚基能结合镁离子,具有β螺旋桨(β-propeller)结构,内有多个叫做WD的重复序列,其功能是在结构上稳定β亚基与α亚基的结合。γ亚基含有异戊二烯化修饰,通常与β亚基形成复合物。Gi蛋白与Gs非常相似,它们含有相同的β和γ亚基,但含有的α亚基不同。Gi也能结合和水解GTP,能被百日咳毒素修饰,但不能被霍乱毒素修饰。

(2) 小G蛋白 一般只有一条肽链组成,在结构上与三聚体G蛋白的α亚基相似。主要包括6类:①Ras蛋白,参与许多生长因子的信号转导(见后);②Ran蛋白,帮助蛋白质进出细胞核;③Rab蛋白,参与真核细胞内的小泡定向和融合;④ARF蛋白,参与形成小泡外被体(vesicle coatomer)包被;⑤Rho蛋白,调节肌动蛋白细胞骨架的组装;⑥参与蛋白质生物合成的某些起始因子、延伸因子和终止因子(参看第三十九章“蛋白质的生物合成及其在细胞内的降解”)。

所有的G蛋白都有两种形式:一种是与GDP结合的无活性形式,另外一种是与GTP结合的活性形式。这两种形式可以相互转变(图17-12),一方面GTP可以取代GDP,另一方面GTP也可以水解成GDP。因此,G蛋白在细胞内相当于是一种分子开关,可通过它在“开和关”之间的切换来控制细胞内的多项活动。细胞内[GTP]/[GDP]一般维持在10:1,这显然有利于GTP取代GDP。

当异源三聚体G蛋白结合GDP时,α亚基与β和γ亚基结合在一起,没有活性。一旦G蛋白结合GTP,α亚基就会与β和γ亚基解离,随后去激活效应器。由β和γ亚基组成的二聚体有时也会起作用。Gα的鸟苷酸结合位点由几个环组成,这些环从由6股β折叠构成的边伸展出来。另外,还有3个转换结构域(switch domain)在GTP取代GDP以后改变位置,这些结构域包括与GTP的γ磷酸基团相邻的残基,以及与β、γ两个磷酸基团结合的Mg^{2+}。

Quiz5 试设计一个实验,证明G蛋白的激活是GTP取代GDP,而不是GDP接受一个磷酸基团变成GTP。

一旦与G蛋白结合的GTP水解,G蛋白就失去活性。GTP的水解是由水分子亲核进攻其末端的磷酸基团引起的。Gα的第二个转换结构域含有一个保守的Gln残基,有助于拉近进攻的水分子,使其靠近与活性中心结合的GTP。

绝大多数G蛋白的功能依赖于两类辅助蛋白:一类是促进GDP/GTP交换的鸟苷酸交换因子(guanine nucleotide exchange factor,GEF),它使G蛋白从与GDP结合的形式变为与GTP结合的形式,

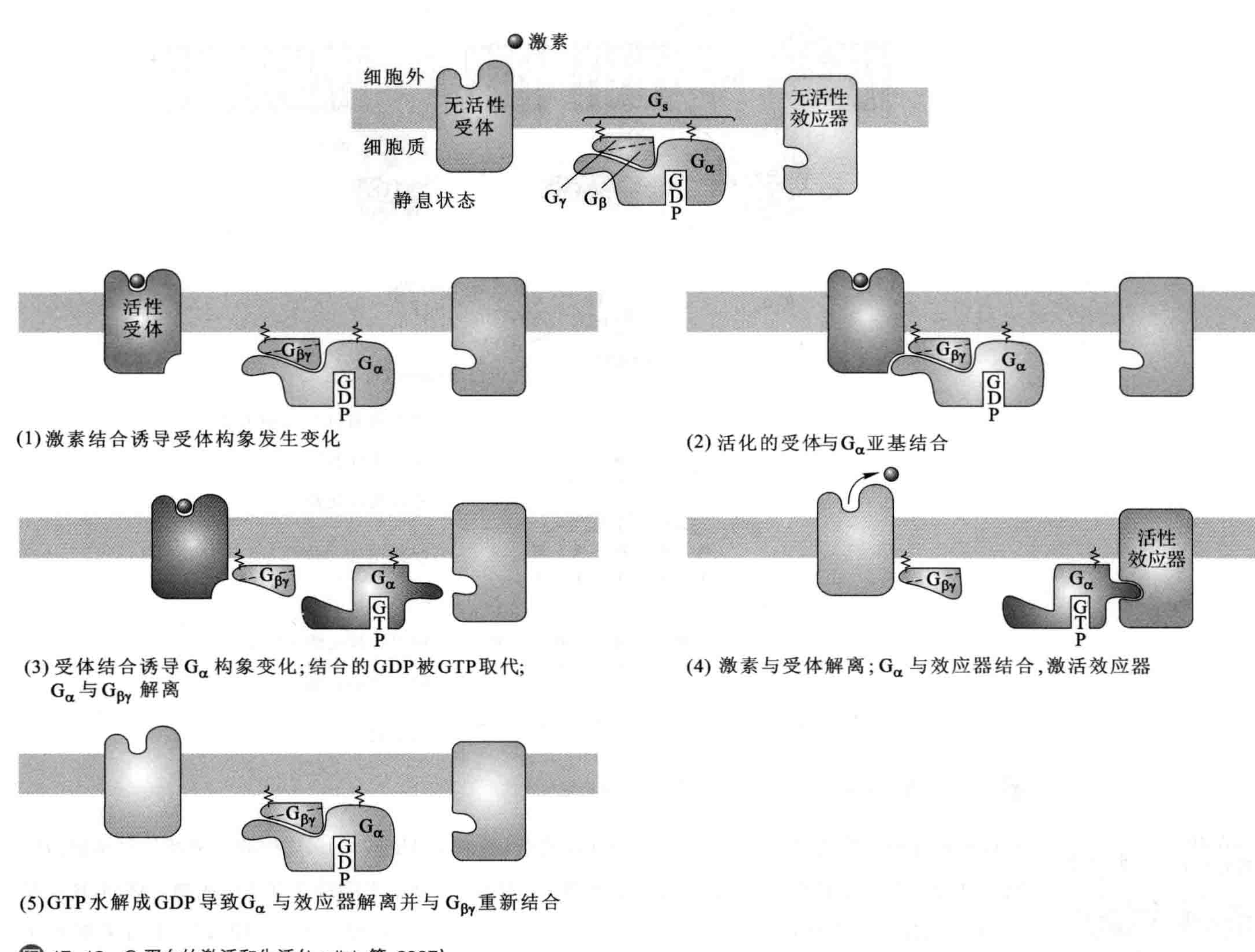


图 17-12 G 蛋白的激活和失活(Lodish 等,2007)

另外一类是促进 GTP 水解的 GTP 酶激活蛋白(GTPase activating protein,GAP)。GAP 是激活小 G 蛋白的 GTP 酶活性所必需的,原因可能是它能插入一个带正电荷的 Arg 残基到活性中心,通过这个 Arg 残基带有的正电荷与 GTP 末端磷酸基团带负电荷的 O 原子的相互作用,来稳定水解过程中形成的过渡态。异源三聚体 G 蛋白的 α 亚基单独就能催化 GTP 的缓慢水解,原因是它已含有必需的 Arg 残基,不需要 GAP 提供,但如果有 GAP,也能提高 α 亚基的 GTP 酶活性。

GPCR 是细胞最常见的一类受体,G 蛋白被这类受体激活以后,要么去激活靶细胞膜上特定的酶的活性,要么去调节特定的离子通道的活性。被激活的 GPCR 实际上就是一种特殊的 GEF。

1. 以 Gs 蛋白偶联的腺苷酸环化酶系统

以肾上腺素为例,其靶细胞不止一种,其受体也不止一种。肝细胞的质膜上有肾上腺素的 β 受体。β 受体是一个典型的 7 次跨膜受体。β 受体与肾上腺素结合以后,构象即发生变化,充当 GEF,刺激细胞质中的 GTP 取代原来与 Gsα 亚基结合的 GDP,Gs 蛋白因此被激活。Gsα-GTP 在与 β 和 γ 亚基解离以后,可独立激活位于细胞膜内侧的效应器即 AC 的活性。激活的 AC 以 ATP 为底物,催化 cAMP 的合成。cAMP 作为肾上腺素的第二信使,再代替肾上腺素去激活信号通路的下一个成分(图 17-13)。

蛋白质激酶 A(protein kinase A,PKA)是一种受 cAMP 直接激活的蛋白质激酶,它催化 ATP 分子上的 *γ*- 磷酸根转移到蛋白质的 Ser 或 Thr 残基的羟基上。磷酸化的 Ser/Thr 通常位于 RRXS/TX 序列(X 为疏水氨基酸残基)模体中。

PKA 由两个调节亚基(R 亚基)和两个催化亚基(C 亚基)组成,但以这种全酶形式存在的 PKA 并无催化活性,原因是 R 亚基上含有一个假底物序列(pseudosubstrate sequence)——RRXAX,假底

Quiz6 如果将 R 亚基上的假底物序列突变成真底物序列,后果会是什么?

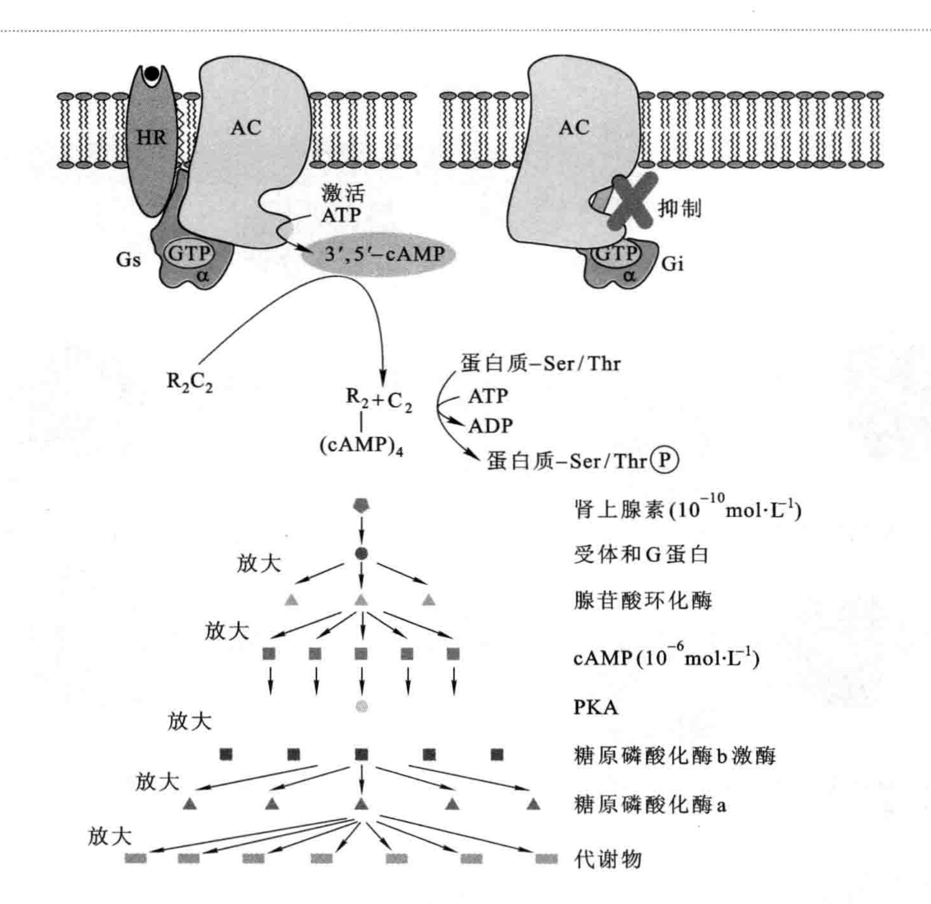


图 17-13 肾上腺素介导的腺苷酸环化酶系统及其级联放大作用

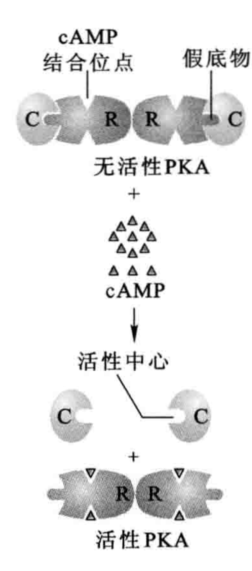


图 17-14 PKA 的激活

物序列假在里面没有真底物序列有的 Ser/Thr 残基(换成了 Ala 残基)。这种缺乏磷酸化羟基的假底物序列与 C 亚基的活性中心结合,等于是封闭了活性中心,使其不能结合真正的底物。然而,R 亚基可结合 cAMP,一个 R 亚基可结合两个 cAMP 分子。一旦 R 亚基结合了 cAMP,即与 C 亚基解离,C 亚基上的活性中心就得以充分暴露,PKA 即被激活(图 17-14)。具体的反应式是:

$$R_2C_2(\text{非活性}) + 4cAMP \longrightarrow (R\cdot cAMP_2)_2 + 2C(\text{活性})$$

每一个 C 亚基都可催化靶蛋白底物序列上的 Ser 或 Thr 残基发生磷酸化修饰。

细胞内有一种 A 激酶锚定蛋白(A-kinase anchoring protein,AKAP),它能与 PKA 的 R 亚基结合,从而将 PKA 限定在细胞特定的区域。一些蛋白质激酶抑制剂(protein kinase inhibitor,PKI)也能通过自带的假底物序列,抑制 C 亚基的活性。

PKA 的底物有许多种,但肝细胞中的主要底物是糖原磷酸化酶 b 激酶。该酶的组成是 $\alpha_4\beta_4\gamma_4\delta_4$,其中 δ 亚基为钙调蛋白(calmodulin,CaM),能够结合钙离子,γ 亚基为催化亚基,α 和 β 亚基为调节亚基,能被磷酸化修饰。一旦受到 PKA 的催化发生磷酸化修饰,糖原磷酸化酶 b 激酶就被激活,并以糖原磷酸化酶 b 为底物,使其发生磷酸化修饰进而转变成有活性的糖原磷酸化酶 a。糖原磷酸化酶 a 即可以催化糖原的磷酸解反应。在整个信号转导过程中,多个环节出现一种酶以另一种酶作为底物而导致信号层层放大的现象。激素作用能够产生强大的生物学效应主要与此有关(图 17-13)。

PKA 除了通过其 C 亚基发挥作用以外,其 R 亚基在特定的情况下也能独立地行使功能。例如,R 亚基能够配合 C 亚基的作用,去抑制一种磷蛋白磷酸酶的活性。这种磷蛋白磷酸酶负责终止由 PKA 激活的糖原磷酸解作用。

PKA 除了能直接调节许多蛋白质或酶的活性而产生“快反应”以外,还能通过激活一种特殊的转录因子来调节特定的基因表达,从而产生“慢反应”。产生“慢反应”的具体过程是(图 17-15):cAMP 作用 PKA,导致 PKA 的 R 亚基和 C 亚基解离。单独的 C 亚基进入细胞核,催化 CRE 结合蛋白(cAMP response element binding protein,CREBP)的磷酸化(Ser 113)。被磷酸化修饰的 CREBP 与 DNA 分子

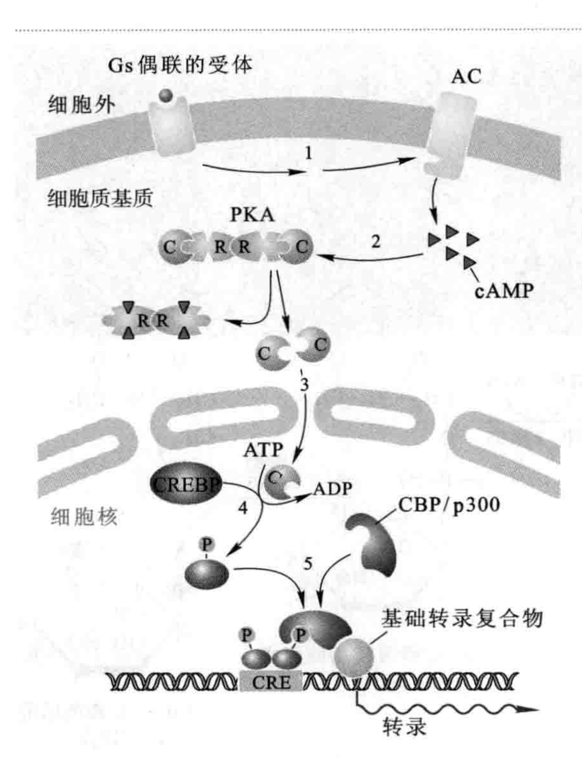


图 17-15 肾上腺素通过 PKA 的"慢反应"(Lodish 等,2007)

上的 cAMP 应答元件(CRE)结合,然后将由 CREBP 结合蛋白(CBP)和 p300 蛋白所形成的 HAT 酶复合物(p300/CBP)招募过来,共同激活下游基因的表达(参看第四十三章"真核生物的基因表达调控")。CRE 的一致序列为 TGACGTCA。

使用 Gi 蛋白的例子是肾上腺素通过 α_2 受体作用胰腺的 β 细胞,致使 β 细胞膜上的 AC 被抑制,进而抑制胰岛素的分泌。不过,Gi 蛋白是如何抑制 AC 尚不十分清楚。一种观点认为,Gi 与 Gs 一样可直接作用 AC,由 Giα–GTP 直接抑制 AC;另一种观点则认为,抑制是由 β 和 γ 亚基完成的,具体的过程是:肾上腺素与 α_2 受体的结合导致了 Gi 蛋白的 α 亚基与 βγ 二聚体的解离,解离出来的 βγ 二聚体与 Gsα 缔合成 Gs 蛋白,从而终止 Gs 蛋白对 AC 的激活。

心肌的收缩同时受到肾上腺素和乙酰胆碱的控制,这两种激素分别通过 Gs 和 Gi 起作用:①与肾上腺素 β 受体偶联的 Gs 能提高 cAMP 浓度,导致心肌的强收缩;②与乙酰胆碱受体偶联的 Gi 蛋白,能降低 cAMP 浓度,导致心肌的弱收缩。从 Gi 释放出来的 β、γ 亚基能够与从 Gs 释放出来的 Gsα 亚基结合,从而减弱肾上腺素的效应。

2. 与 Gq 蛋白偶联的磷酸肌醇系统

另一个典型的 G 蛋白系统是磷酸肌醇系统(图 17-16)。现以下丘脑分泌的促性腺激素释放激素(gonandotropin-releasing hormone,GnRH)为例加以说明。GnRH 的受体位于某些脑垂体前叶细胞上,当 GnRH 与其受体结合以后,受体的构象发生变化,激素与受体的二元复合物激活 Gq。

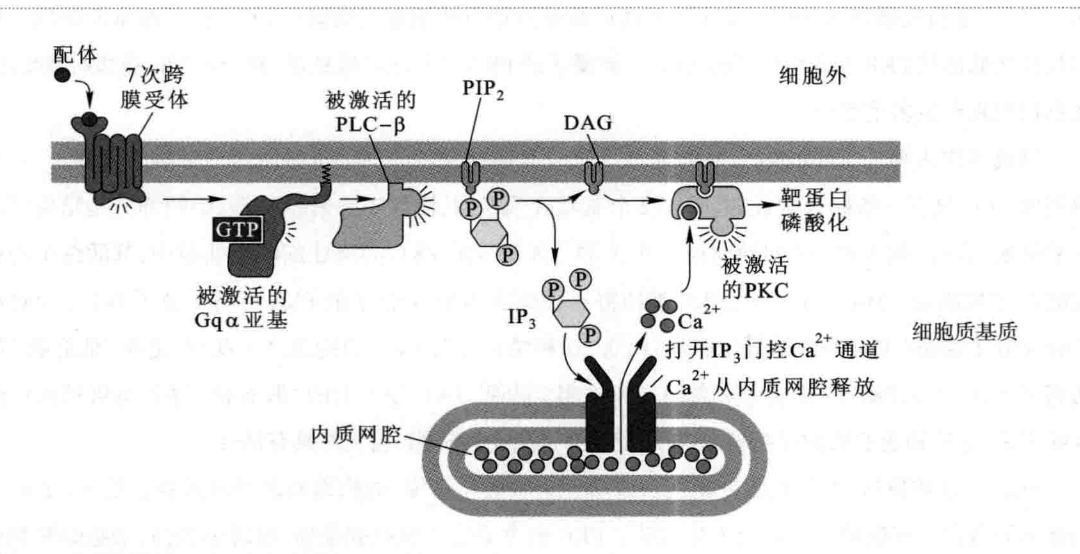


图 17-16 磷酸肌醇系统的详细图解(Alberts 等,2007)

激活的机制类似于 Gs 的激活,但与 Gs 不同的是,受 Gq 激活的效应器是一种对磷脂酰肌醇特异性的磷脂酶 C(PLC-γ)。这种 PLC 正常的底物是位于细胞膜内侧的磷脂酰 4,5- 二磷酸肌醇(PIP_2)(图 17-17),PIP_2 由此被水解为 DG 和 IP_3,这两种小分子物质为 GnRH 的第二信使。其中 DG 脂溶性较强,并不离开细胞膜,而 IP_3 是一种可溶性的小分子,生成后即被释放到细胞质。内质网膜上有 IP_3 的受体,在 IP_3 与其结合后可促使周围的 Ca^{2+} 通道开放,于是贮存在内质网腔中的 Ca^{2+} 得以释放,细胞质 Ca^{2+}

图 17-17 磷脂酰肌醇的代谢转变

浓度便迅速升高。Ca^{2+} 可视为另一种第二信使，也可以视为一种第三信使，它可影响到细胞内多种蛋白质或酶的活性，调节许多重要的生理功能，如肌肉收缩、细胞周期控制和细胞分化。

IP_3 还可以进一步磷酸化成 IP_4、IP_5 或 IP_6。这几种磷酸化的肌醇也可以充当信使分子。例如，1,3,4,5- 四磷酸肌醇(inositol-1,3,4,5-tetraphosphate,IP_4)在某些细胞可激活质膜上的 Ca^{2+} 通道，使胞外的 Ca^{2+} 也能涌入胞内。此外，PI 的各种磷酸化产物——PI-3-P、PI-3,4-P_2、PI-3,4,5-P_3 和 PI-4,5-P_2 也具有特殊的信号作用。它们作为特定的血小板 - 白细胞 C 激酶底物同源(pleckstrin homology,PH)结构域和 FYVE 蛋白结构域的配体，有助于将含有这两种结构域的蛋白质招募到膜表面。例如，蛋白质激酶 B(PKB)以这种方式从细胞质基质被招募到质膜上以后，位于细胞质基质一侧的其他激酶催化 PKB 的磷酸化，使之激活。被激活的 PKB 再催化其他靶蛋白的 Ser/Thr 残基的磷酸化，诱发靶细胞产生多重效应。

钙离子作为第二信使的第一个功能，是与 DG 和磷脂酰丝氨酸一道激活蛋白激酶 C(PKC)。一个典型的 PKC 只有一条肽链构成，共有 672 个氨基酸残基，其活性中心紧靠 C 端，中间部分是结合钙离子的区域，含有“假底物”序列的调控区在 N 端。无活性的 PKC 溶解在细胞质基质中，其活性中心被假底物序列遮盖。DG 产生以后，PKC 被招募到细胞膜内侧，1 分子的 PKC 可与 1 分子 DG、1 个钙离子和 4 分子磷脂(主要是磷脂酰丝氨酸)结合，这种结合导致 PKC 的构象发生变化，迫使“假底物”序列离开活性中心，PKC 因此被彻底激活。体外实验表明，PKC 分子上的“假底物”序列也可被蛋白酶水解下来，这种缺乏假底物序列的 PKC 即使没有钙离子和磷脂，也依然具有活性。

PKC 一旦被激活，可催化多种靶蛋白或酶发生磷酸化修饰，被修饰的氨基酸残基也是 Ser/Thr，受到修饰的蛋白质或酶的活性将会发生变化。PKC 的主要底物包括膜受体、膜转运蛋白、细胞骨架和许多重要代谢途径中的限速酶。其中对细胞骨架的作用与细胞的分泌相关，GnRH 促进脑垂体前叶分泌促性腺激素就是这种功能的具体表现。

钙离子传感器蛋白(sensor)是另一类受到钙离子激活的蛋白质。此类蛋白都具有 α 螺旋 - 环 -α 螺旋或 EF 手相这种结构模体(参看第二章“蛋白质的结构”)，其中环能结合钙离子。在较低的 Ca^{2+} 浓度下(10^{-8}~10^{-7} mol/L)，各种 Ca^{2+} 传感器蛋白无活性；当 Ca^{2+} 浓度提高到 10^{-6}~10^{-5} mol/L，它们的活性才能表现出来。

由美籍华裔科学家张槐耀首先发现的钙调蛋白(CaM)就是其中最重要的一种。CaM是一种对热稳定的酸性蛋白,由148个氨基酸残基组成,M_r为1.7×10^3,它存在于所有的真核细胞,在进化上具有高度的保守性,如酵母细胞与人的CaM在一级结构上仅相差3个氨基酸。这种蛋白质在三维结构上像一个哑铃,一段7圈长的α螺旋将两个球叶相连,每一个球叶具有两个α螺旋-环-α螺旋,每一个α螺旋-环-α螺旋能结合一个钙离子。两个球叶之间的α螺旋的N端呈疏水性,C端呈酸性。这段α螺旋(主要是酸性部分)就是CaM与其他蛋白质或酶(两亲螺旋的碱性面)相互作用的区域。

对CaM的激活机制的研究已有突破性的进展,许多证据表明,在没有钙离子的情况下,CaM的两球叶塌陷,环绕在球叶之间的α螺旋上,α螺旋上的疏水段和酸性区段因此被遮盖,这时CaM没有任何活性;一旦结合钙离子,CaM的构象立刻发生变化,两球叶发生移动,球叶之间的螺旋被暴露,相关的蛋白或酶就能与它发生相互作用,它们的活性也因此被激活。

能受到CaM调节的蛋白质或酶很多,如cAMP磷酸二酯酶和肌球蛋白轻链激酶。在这些例子中,CaM与相应的酶结合,临时作为这些酶的亚基,以调节它们的活性。在糖原磷酸化酶b激酶中,它则是作为一个永久性亚基发挥调节的作用。

受CaM激活的另一个典型代表是依赖于CaM的蛋白质激酶(CaM dependent protein kinase,CaMPK)。该激酶由12个亚基组成,这些亚基排列成双层六角形。构成激酶的亚基有α和β,这两种亚基在结构上高度同源,可以按照不同的比例结合在一起而形成不同形式的同工酶。在脑细胞中,两者的比例是9∶3。所有的亚基都能够与CaM结合,具有催化活性,并能发生自我磷酸化。CaM的作用是刺激它的激酶活性,激酶的活性进而可催化自身的磷酸化反应,而这种自我磷酸化作用使得激酶的活性不再依赖于CaM和钙离子。在脑细胞中,这种蛋白质激酶的含量非常丰富,它最合适的底物是细胞骨架。

钙调磷酸酶(calcineurin)是一种分布广泛的多功能磷酸酶,也受到CaM的调节。这种磷酸酶可水解磷蛋白分子上的磷酸基团,从而抵消激酶的作用。这种酶也由α和β两个亚基组成,其中α亚基能结合锌离子和铁离子,并含有活性中心和CaM结合位点,β亚基有4个钙离子结合位点。

通过Gq蛋白作用的激素除了GnRH,还有TRF、血管紧张素Ⅱ/Ⅲ、肾上腺素(通过α_1受体)和AVP(作用脑垂体前叶细胞协助ACTH的分泌)等。

在磷酸肌醇系统中,Ca^{2+}对PKC的激活与Ca^{2+}通过CaM对CaMPK的激活之间常常是相互协调,或在时间上密切配合。

3. G蛋白偶联的受体作用系统的终止

从以上各种G蛋白系统作用的具体过程可以看出,G蛋白系统一般主要通过激素→GPCR→G蛋白→效应器→第二信使→蛋白质激酶→靶酶或靶蛋白这样的途径依次进行信号转导,并最终诱发靶细胞产生特殊的生物学效应。与其他系统一样,G蛋白系统在靶细胞中不能始终处于活化状态。细胞必须存在一套有效的灭活机制,能使靶细胞重新恢复到最初的静息状态,以便接受下一轮刺激。

有以下几种机制与G蛋白系统的灭活有关:

(1) HR解离　激素停止分泌,其浓度的下降可导致HR的解离,而一旦HR解离,受体就恢复到非活性构象状态。解离下来的配体可能被代谢分解掉或者被重吸收。例如,乙酰胆碱被乙酰胆碱酯酶水解成胆碱和乙酸,肾上腺素被单胺氧化酶氧化成无活性的醛类化合物或者被分泌它的细胞重新吸收。

(2) 受体脱敏　这种过程会因激素的不同而不同。某些受体经G蛋白受体激酶(G-protein receptor kinase,GPRK)的作用发生磷酸化修饰。磷酸化的受体能与胞内的*β*-拘留蛋白(arrestin)结合。这种蛋白质一方面抑制了受体激活G蛋白的活性,另一方面充当网格蛋白的适配体,使受体更容易发生由网格蛋白介导的内吞,从而使质膜上的受体数目下降。

(3) 第二信使的降解或去除　G蛋白系统的运转与第二信使的作用是分不开的,但是无论哪一种第二信使在细胞内都有一定的半衰期,比如cAMP可被A-PDE(一种专门水解cAMP的磷酸二酯酶)催化水解为5′-AMP。巧妙的是,A-PDE本身就是PKA的一个底物,在PKA激活以后不久,A-PDE也

很快被磷酸化而激活去水解 cAMP。当 cAMP 被水解后，PKA 将回到四聚体状态，PKA 的活性由此丧失。又如磷酸肌醇系统中产生的 IP_3 可被磷酸酶依次水解成 IP_2、IP、I，I 可作为重新合成 PI 的原料，PI 在激酶的催化下形成 PIP_2，DG 则可通过脂循环重新转变为磷脂，或者被水解酶进一步水解成甘油和脂肪酸，其中从 DG 的 2 号位羟基水解下来的脂肪酸一般是花生四烯酸，花生四烯酸可转变为前列腺素等许多生理活性物质，作用于临近的细胞，引起其他的生物学效应。

躁狂症是一种常见的精神疾患，患者脑细胞上的磷酸肌醇系统过于活跃。Li^+ 可被用来治疗躁狂症，原因是它能在体内阻止 IP→I，这也就阻止了 PIP_2 的再生，从某种程度上也就抑制了 IP_3 和 DG 的形成。

钙离子可被钙泵重新泵回内质网腔，这使那些依赖于钙离子激活的蛋白质和酶恢复到原来的无活性状态。

第二信使的代谢若受到抑制，无疑会延长激素的作用时间。例如，茶碱和咖啡因能够抑制 A-PDE 的活性，因此，这两种物质在体内能够延长肾上腺素的作用时间，即有兴奋机体的功能。佛波脂（phorbol ester）是一种天然的肿瘤促进剂，原因是其分子结构的一部分和 DG 非常相似（图 17-18），在体内能够直接激活 PKC。但由于佛波脂自身很难被代谢掉，PKC 将受到它的持续激活，最终可导致细胞的癌变。

图 17-18　佛波脂的化学结构

（4）G 蛋白的自我灭活　这是由 G 蛋白本身具有的 GTP 酶活性来完成的。已知 G 蛋白的 α 亚基具有 GTP 酶活性，虽然活性不高，但是仍然可使 G 蛋白从有活性的状态变成无活性的状态。G 蛋白所具有的 GTP 酶活性不可以很高，否则 G 蛋白还没来得及起作用就把自己灭活掉了。

许多 G 蛋白的 GTP 酶活性还需要 GAP 激活。GAP 能提高 GTP 酶活性，从而加快 G 蛋白灭活的速率。如果 G 蛋白的 GTP 酶活性受到抑制，将会导致靶细胞持续处于活化状态，而影响细胞的正常功能。McCune Albright 综合征是一种骨骼异常和骨囊肿疾患，其病因是 Gsα 亚基的 Arg201 突变为 His201 或 Cys201，从而导致 GTP 酶失活，引起 AC 持续的激活。

此外，霍乱的发病机理也与此相关。霍乱是一种肠道传染病，由霍乱弧菌引起。霍乱弧菌可产生一种外毒素即霍乱毒素（cholera toxin，CT），CT 由 1 个 A1 亚基、1 个 A2 亚基和 4 个 B 亚基组成，A1 亚基和 A2 亚基之间通过二硫键相连，在小肠上皮细胞的细胞膜上有它的受体。CT 作用小肠上皮细胞的第一步是 B 亚基与受体结合，随后 A1 亚基和 A2 亚基之间的二硫键被还原，A1 亚基随即进入细胞内。在细胞内，A1 亚基作为一种特殊的转移酶，催化 NAD^+ 分子上的 ADP- 核糖基转移到 Gsα 亚基的 Arg201 残基上（与 McCune Albright 综合征突变的氨基酸残基相同），反应式为：

Quiz7 你还记得霍乱毒素受体的化学本质吗？

$$\text{Gs-Arg} + \text{NAD}^+ \longrightarrow \text{Gs-Arg-核糖-ADP} + \text{烟酰胺}$$

Gs 蛋白被修饰以后，其 GTP 酶的活性由此丧失。由于小肠对水分的排泄依赖于 Gs 蛋白对 AC 的激活，因此 Gs 蛋白的 GTP 酶活性的丧失必然会导致 AC 的持续活化。这使得 cAMP 大量被合成，小肠便在短时间内丢失大量水分和电解质。患者因此出现持续的腹泻和呕吐症状，严重可致死。

GTP 的类似物 GTPNH 和 $GTPCH_2$（分别由 N 和 C 取代 GTP 分子上 β 和 γ 磷酸根之间的 O），能够代替 GTP 与 G 蛋白的 α 亚基结合，但却不能被 GTP 酶水解，因此，在细胞中加入这几种 GTP 的类似物也可导致 G 蛋白系统的持续活化。

百日咳（pertussis，whooping cough）是由百日咳杆菌所致的急性呼吸道传染病。这种细菌能产生一种外毒素——百日咳毒素（pertussis toxin，PT）。PT 由 1 个 A 亚基和 5 个 B 亚基组成。A 亚基能进入呼吸道上皮细胞，作为一种转移酶来催化 NAD^+ 上的 ADP- 核糖基转移到 Giα 亚基的一个 Cys 残基上，这种修饰可阻止结合在 Giα 上的 GDP 与细胞质基质中的 GTP 进行交换，也就阻止了 Gi 蛋白抑制 AC。

Quiz8 CT 和 PT 的作用是否都可导致各自靶细胞内的 cAMP 水平升高？

（5）蛋白质的去磷酸化　除了嗅觉、味觉和视觉以外，G 蛋白系统最终一般都会涉及一些蛋白质

激酶(如 PKA、PKC 和依赖于 CaM 的蛋白质激酶)对靶蛋白进行磷酸化修饰。然而,细胞内还存在另一类催化磷酸根水解的磷蛋白磷酸酶(phosphorylated protein phosphatase,PPP)。PPP 所起的作用是让被激酶修饰的蛋白质恢复到原来的脱磷酸化状态。

除了多种激素可通过 G 蛋白进行信号转导以外,高等动物视觉、嗅觉以及味觉的形成虽然不是由激素介导的,但其中也涉及不同的 G 蛋白。其中视觉的产生需要 Gt 蛋白,嗅觉的产生需要 Golf 蛋白,味觉的产生需要味觉素。

e17-2 视觉、嗅觉和味觉产生的分子机制

（二）酶受体系统

1. 受体鸟苷酸环化酶系统

该系统的受体具有潜在的鸟苷酸环化酶(GC)活性,不需要 G 蛋白,以 cGMP 作为第二信使。通过此系统起作用的激素有:心房利钠肽(atrial natriuretic peptide,ANP)、脑利钠肽(brain natriuretic peptide,BNP)和 C 型利钠肽(C-type natriuretic peptide,CNP)。

ANP 也称为心钠素或心房利钠因子(atrial natriuretic factor,ANF),由心房上特殊的细胞合成和分泌。PKC 能促进它的分泌,PKA 则能抑制它的分泌。ANP 主要作用肾小管细胞,其受体就横跨在这种细胞的质膜上,只由一条肽链组成,在细胞膜的外侧含有 ANP 结合部位,在细胞质一侧具有 GC 活性部位。

如图 17-19 所示,ANP 的结合会诱导受体构象发生变化,受体位于细胞质一侧的 GC 随后被激活。GC 催化细胞质里的 GTP 环化形成 cGMP,后者作为第二信使,激活依赖于 cGMP 的蛋白质激酶(cGMP-dependent protein kinase,PKG)。PKG 由 2 个完全相同的亚基组成,每一个亚基的 M_r 为 7.6×10^4,其中前 100 个氨基酸残基负责亚基之间的聚合,在 C 端含有激酶的活性中心。每一个亚基能结合 2 分子的 cGMP,结合 cGMP 的部位结构与 PKA 的 R 亚基结合 cAMP 的部位结构同源。当每一个亚基各结合 2 个 cGMP 以后,PKG 的构象即发生变化,其蛋白激酶的活性被激活。与 PKA、PKC 一样的是,PKG 也是一种 Ser/Thr 激酶,被激活后可催化靶细胞内的一系列靶蛋白或靶酶进行磷酸化修饰,而产生特殊的生物学效应。

2. 受体酪氨酸蛋白激酶系统(receptor tyrosine kinase,RTK)

该系统对于细胞的生长和分裂极为重要,具有以下几个比较重要的特征:

(1) 通过该系统发挥作用的激素主要是一些生长因子 例如:胰岛素、表皮生长因子(epidermal growth factor,EGF)、血小板衍生的生长因子(platelet derived growth factor,PDGF)、集落刺激因子 -1(colony stimulating factor-1,CSF-1)、神经生长因子(nerve growth factor,NGF)以及类胰岛素生长因子 -Ⅰ和Ⅱ(insulin like growth factor,IGF)等。生长因子是一类能够促进细胞生长和分裂的物质,在化学本质上通常为多肽或蛋白质,一般以旁分泌或者自分泌的方式起作用。

(2) 受体具有潜在的酪氨酸蛋白激酶的活性

(3) 受体具有高度保守的结构,一般可分成 3 个功能区:配体结合区、跨膜区和胞内酪氨酸激酶区(图 17-20)。

配体结合区面向胞外,位于肽链的 N 端,是与生长因子结合的区域。该区域具有多个潜在的糖基化位点(Asn-X-Ser/Thr,X 代表任何氨基酸),且富含 Cys 残基;跨膜区是一段单跨膜的疏水 α 螺旋,它将配体结合区和激酶活性区联系在一起;酪氨酸激酶区的活性受相关的生长因子的控制。当与相应的生长因子结合以后,受体的构象发生变化,酪氨酸激酶区的活性可提高 3 倍之多。激酶的底物包括受体本身以及细胞内其他一些蛋白质。激酶的底物若是受体本身,就导致受体的自我磷酸化修饰。这种自我磷酸化修饰改变了受体的电荷和构象,不仅能够进一步刺激受体所具有的酪氨酸激酶的活性,还为招募含有 SH2(Src-homology 2)结构域的蛋白质(例如 Src、Shc、CB1、GAP、PLCγ、磷脂酰肌醇激酶和 Grb2/Sem-5 蛋白)与受体的结合创造了条件(图 17-21),由此也启动了多种蛋白质之间的相互作用和磷酸化的级联反应。

(4) 一般都会激活特定基因的表达。这是将胞外信号转导到细胞核最重要的途径。

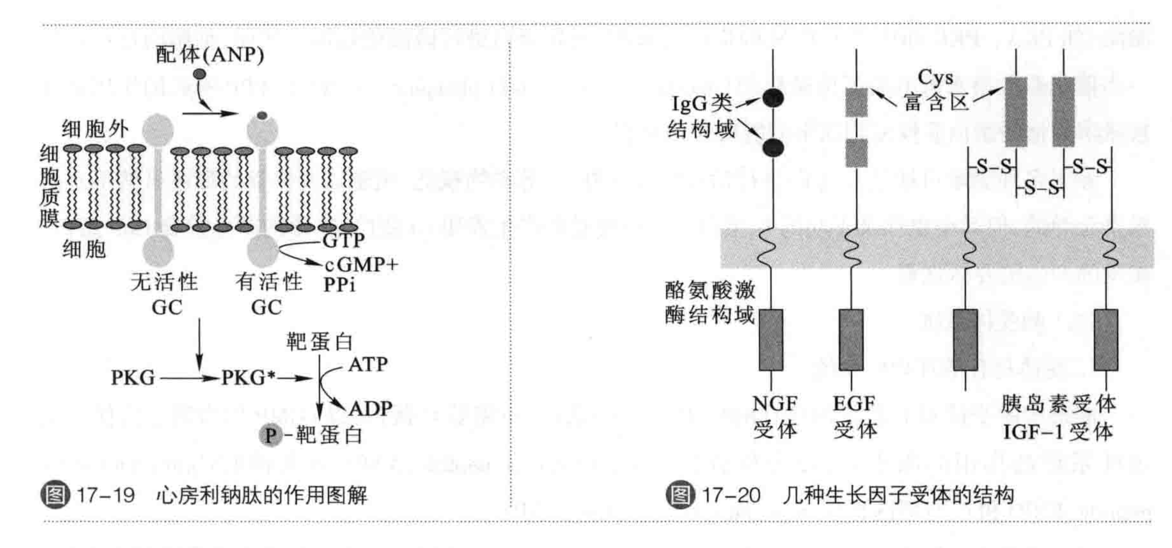


图 17-19　心房利钠肽的作用图解

图 17-20　几种生长因子受体的结构

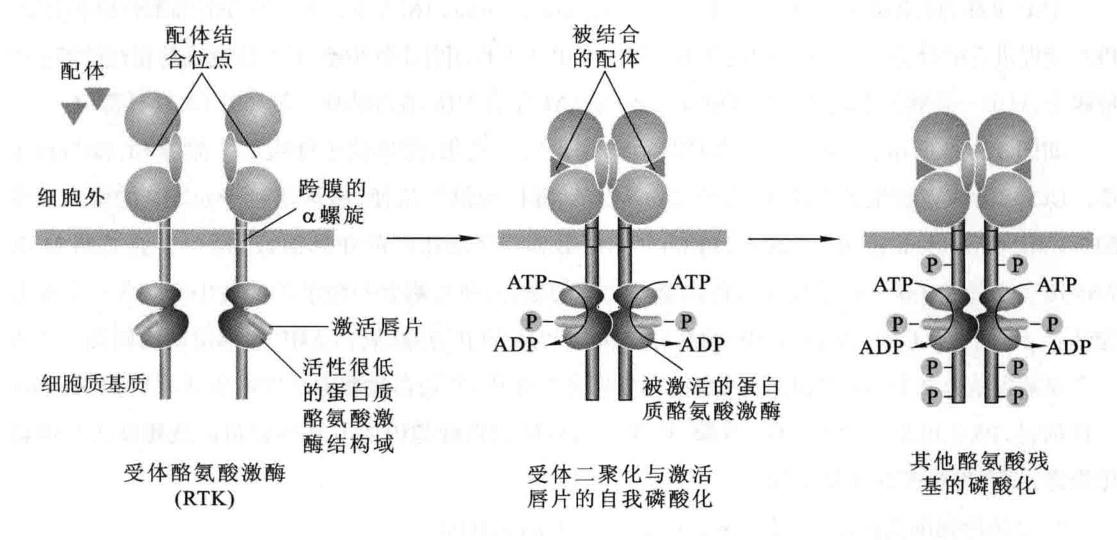


图 17-21　RTK 系统的受体激活

(5) 酪氨酸残基的脱磷酸化由专门的蛋白质酪氨酸磷酸酶(protein tyrosine phosphatase, PTP)催化。磷酸酶的作用是逆转由激酶引发的反应,其中某些磷酸酶也作为受体(如 CD45 抗原)定位在细胞膜上。

(6) 与细胞的癌变有密切的联系　细胞的癌变是指细胞的分裂失去控制,而引起细胞癌变的因素较为复杂,有理化因素(如紫外线和黄曲霉素),还有生物学因素(如肿瘤病毒)。分子水平的研究已表明,在真核生物和肿瘤病毒的基因组内有一类与癌症发生相关的基因,这类基因被称为癌基因(oncogene)。其中存在于真核生物基因组中的癌基因被称为细胞癌基因或原癌基因(cellular oncogene, c-onc),而存在于肿瘤病毒基因组之中的癌基因被称为病毒癌基因(viral oncogene, v-onc)。原癌基因在正常的细胞内是良性的,其表达产物为细胞的功能所必需,主要起着调节细胞生长和分化的作用。当它们受到各种因素的作用而发生突变,或者在细胞中过量表达之后,可导致细胞的癌变。

现已发现有近百种结构不同的原癌基因,根据表达产物的化学本质和生物功能特征,大致可将它们划分为以下几类:①蛋白质激酶类,这类癌基因的产物具有蛋白质激酶的活性,有的具有酪氨酸激酶的活性,有的具有丝氨酸 / 苏氨酸激酶活性(如 c-Src 和 c-Fps);②生长因子类,这类癌基因的产物本质上也属于生长因子(例如 c-Sis);③生长因子受体类,这类癌基因的产物定位于某些细胞的质膜上,充当生长因子的受体(例如 c-Erb B);④ G 蛋白类(如 c-Ras);⑤转录因子类,这类癌基因的产物是一类调节基因转录的蛋白质因子(例如 c-Fos 和 c-Jun)。

病毒癌基因则随着病毒感染细胞而进入胞内,在利用宿主细胞的表达系统得到表达后,可直接作用宿主细胞内与细胞生长和分裂有关的信号通路,干扰正常的信号转导而导致细胞癌变。与原癌基

因相比，病毒癌基因的产物大多数是某种生长因子和某些激素作用系统中某一成分的类似物。这些病毒癌基因在正常的细胞中表达以后，持续模拟生长因子或激素的效应，而导致细胞的癌变。例如猿猴肉瘤病毒的癌基因 *v-sis* 在受宿主细胞中表达的产物与 PDGF 的一个亚基几乎是相同的，因此受猿猴肉瘤病毒感染的细胞会大量表达这种 PDGF 的类似物。大量的 PDGF 类似物可而持续地刺激宿主细胞的生长和分裂，而导致其发生癌变。再如一种名为 *v-erbB* 的病毒癌基因，编码的蛋白质是一个截断了的 EGF 受体——缺乏 EGF 的结合部位，但保留着跨膜区和胞内的酪氨酸激酶区。这样的受体类似物不需要结合 EGF 就有了酪氨酸激酶的活性。在已发现的 20 多种逆转录病毒中，有近一半的病毒癌基因所编码的蛋白质具有酪氨酸激酶的活性。

下面分别以 EGF 和胰岛素为例，详细说明 RTK 系统是如何运转的。

EGF 是最早发现的一种多肽生长因子，由动物的唾液腺合成。1972 年，Cohen 首先发现它能够诱导新出生的小鼠眼睑的张开和长牙。进一步研究表明，EGF 能够与血清一起，刺激由外胚层和中胚层衍生而来的细胞的分裂，并可抑制胃液的分泌。

EGF 的受体是第一个被克隆的受体酪氨酸激酶，是单一的糖蛋白。如图 17-22 所示，当与 EGF 结合后，受体在细胞膜上的流动性增强，相互靠近形成二聚体。受体的构象因此发生变化，使其潜在的酪氨酸激酶的活性得以激活，于是单体之间可相互催化对方的磷酸化反应。这种自我磷酸化反应可使激酶的活性进一步增强。当受体分子胞内结构域处于特定位置的酪氨酸残基被磷酸化修饰以后，同时含有 SH2 和 SH3 结构域的接头蛋白（adaptor protein），如 Grb2/Sem-5（Grb2 为哺乳动物细胞中的接头蛋白，Sem-5 为线虫细胞中的接头蛋白），会在一头用它的 SH2 结构域内一个带正电荷的口袋与磷酸化的酪氨酸残基结合，另一头用 SH3 结构域内一个疏水的口袋与 Sos 蛋白的 C 端一段富含 Pro 的疏水肽段结合。于是 Sos 蛋白被招募到细胞质膜，直接作用位于细胞质膜胞质基质一侧的 Ras 蛋白。

Quiz9 一些生长因子受体的抗体能够模拟生长因子的作用，为什么？

图 17-22　RTK 系统的详细图解

在某些细胞中，Grb2/Sem-5·Sos 复合物通过另一种接头蛋白 Shc 与激活的受体结合。Ras 蛋白作为小 G 蛋白，M_r 为 2.1×10^4，通过异戊二烯基团锚定在细胞质膜上。与其他 G 蛋白一样，Ras 蛋白也有两种形式，一种是与 GDP 结合的无活性的形式，另一种是与 GTP 结合的有活性的形式，两种形式可相互转变。

与 Grb2 结合的 Sos 蛋白也是一种鸟苷酸交换因子，它可以促进细胞质基质中的 GTP 取代原来与 Ras 蛋白结合的 GDP，Ras 蛋白因而被激活。被激活的 Ras 蛋白（Ras·GTP）再激活下游的 Ser/Thr 蛋白激酶链，并最终将信号转导到细胞核内，促进细胞的分裂和生长。而 Ras 蛋白所具有的 GTP 酶活性

很快在 GAP 的激活下，催化结合的 GTP 水解成 GDP，从而使 Ras 蛋白恢复到无活性的 Ras-GDP 状态。医学研究表明，有近 30% 的人类癌症含有 *ras* 基因的突变，这种突变绝大多数会导致 GTP 酶活性的丧失。

受 Ras 蛋白激活的 Ser/Thr 蛋白激酶多数是原癌基因的产物，主要包括以下几种：①Raf 蛋白，也称作有丝分裂原激活蛋白激酶的激酶的激酶（mitogen-activated protein kinase kinase kinase，MAPKKK），它由 Ras 蛋白直接激活，也可以受 PKC 催化的磷酸化修饰而激活。当 Raf 蛋白被激活后，可催化有丝分裂原激活蛋白激酶的激酶（mitogen-activated protein kinase kinase，MAPKK）的 Ser/Thr 残基进行磷酸化修饰，从而激活 MAPKK 的活性。②MAPKK 也称为 MEK，是一种双功能激酶，被 MAPKKK 激活化后能对 MAPK 在 TEY（Thr-Glu-Tyr）序列中的 Thr 或 Tyr 残基进行磷酸化修饰，从而激活有丝分裂原激活蛋白激酶（mitogen-activated protein kinase，MAPK）的活性。③MAPK 在受 MAPKK 的激活后，从细胞质基质迁移到细胞核，在细胞核内催化 E1K-1、c-Fos 和 c-Jun 等转录因子的磷酸化，这些转录因子再进一步激活参与细胞分裂的基因表达，如周期蛋白 D1，并最终导致靶细胞的生长和分裂。④MAPK 还能促进 MAPK 磷酸酶 -1（MAPK phosphatase-1，MKP-1）的基因表达，MKP-1 表达以后可催化 MAPK 的去磷酸化，从而终止 MAPK 的作用。

胰岛素的受体由 2 个 α 亚基和 2 个 β 亚基组成，亚基之间有二硫键。胰岛素的结合位点在 α 亚基上。胰岛素与 α 亚基的结合导致 β 亚基的构象发生变化。此时，位于 β 亚基 C 端的酪氨酸激酶结构域催化受体发生自我磷酸化，进一步激活其酪氨酸激酶活性。

胰岛素受体激酶的底物已被发现有 12 种之多，它们统称为 IRS（insulin receptor substrate）。IRS 被磷酸化后，可招募和激活细胞内一系列含有 SH2 结构域的下游蛋白质，例如 PLC-γ、PI3K 和 Grb2/Sos 等，从而进一步激活其他的信号通路。由于不同的细胞含有不同的具有 SH2 结构域的蛋白质，因此，即使配体相同，产生的反应也不一定相同。如果 IRS 激活的是 PLC-γ，磷酸肌醇系统就被启动；如果 IRS 激活 Grb2/Sos，就与 EGF 一样，启动的是 MAPK 通路；如果激活的是 PI3K，细胞内的 PKB 通路就被启动（图 17-23）。PKB 被激活后，不仅可促进葡萄糖转运蛋白 4（GLUT4）从细胞内的小囊泡膜移位到质膜上，以提高葡萄糖的转运（参看第二十一章“生物大分子的消化和吸收”），还能提高蛋白质的合成以及通过磷酸化使糖原合酶抑制剂 GSK3 失活，进而刺激糖原合成。

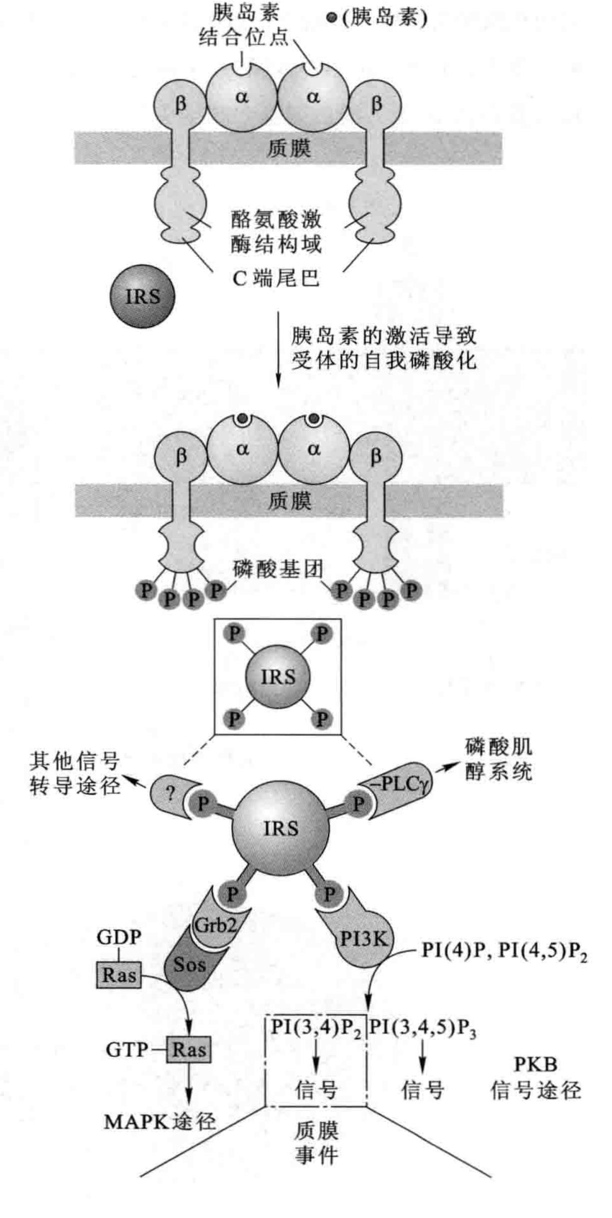


图 17-23 胰岛素作用的分子机制

除了受体鸟苷酸环化酶系统和受体酪氨酸激酶系统之外，受体丝氨酸 / 苏氨酸激酶系统和受体酪氨酸磷酸酶系统也具有潜在酶的活性。这两个系统中受体的酶活性同样只有在相应激素结合上以后才会被激活。

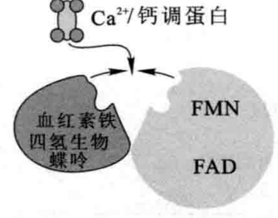
e17-3 受体丝氨酸 / 苏氨酸激酶系统和受体酪氨酸磷酸酶系统

（三）NO 系统

NO 是一种化学性质活泼、具有毒性的小分

子气体，早在1772年就由Joseph Priestly发现，长期以来一直被视为由化石燃料燃烧、汽车尾气和吸烟产生的有毒污染物，它在体内居然能够作为细胞之间的信号分子，这的确大大出乎人们的意料。Ferid Murad在1977年发现，NO气泡通过含有鸟苷酸环化酶(guanylyl cyclase，GC)的组织，可导致组织内cGMP水平的升高，而硝化甘油也能激活相同的GC活性，由此他推测硝化甘油在体内能够释放出NO。1980年Robert F. Furchgott发现，血管内皮细胞在乙酰胆碱的作用下，能产生一种促进血管松弛的因子(endothelium-derived relaxation factor，EDRF)，但他并不能确定EDRF的化学本质；1986年Louis Ignarro也发现，血红蛋白接触到受刺激的内皮细胞产生的EDRF，与接触到NO以后发生的光谱迁移完全相同，这就证明了EDRF就是NO。

NO在体内可介导多种生理功能，特别可在心血管系统、中枢和周围神经系统以及宿主防御等系统中发挥重要作用，并可能参与多种疾病的发生过程。*Science*在1992年曾将其评为当年的"年度分子"，而对该气体的发现和研究有杰出贡献的三位美国科学家——Furchgott、Ignarro和Murad，也因此分享了1998年的诺贝尔生理学或医学奖。



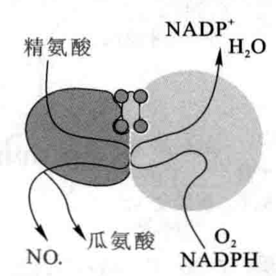


图17–24 NO合酶的结构、辅助因子和总反应式

在体内，NO由一氧化氮合酶(nitric oxide synthase，NOS)催化产生，其主要前体是L–Arg。NOS由两个相同的亚基组成，每个亚基含有1分子FMN、1分子FAD、1分子四氢生物蝶呤和1分子高铁血红素，这些辅助因子的作用是促进Arg失去5个电子而生成NO(图17–24)。已发现有三类NOS：一类是神经元型(neuronal NOS，nNOS)，一类是内皮细胞型(endothelial NOS，eNOS)，还有一类是诱导型(inducible NOS，iNOS)。这三类NOS存在于多种类型的细胞和组织之中，其中nNOS和eNOS主要分别存在于脑细胞和内皮细胞，这两类NOS呈组成型表达，在细胞内Ca^{2+}浓度提高的情况下，会受到Ca^{2+}–CaM的激活。而iNOS主要存在于巨噬细胞、亲核细胞、内皮细胞和平滑肌细胞，但只有在机体处在胁迫的条件下(如细胞受到细胞因子或细菌内毒素的刺激)，它才被诱导表达，而一旦被诱导表达，可存留很长时间。iNOS的活性与Ca^{2+}浓度无关，但仍然依赖于CaM的结合，且与CaM的结合更为紧密，也因此导致体内产生更高水平的NO。

NO的反应性极强，在体内可参与许多反应，很容易通过细胞膜扩散到邻近的细胞，对邻近细胞产生作用。在红细胞内，它能够与氧合血红蛋白或非氧合血红蛋白发生下面的反应：$NO+HbO_2 \longrightarrow NO^{3-}+Hb$或$NO+Hb \longrightarrow Hb(NO)$；在巨噬细胞和亲核细胞中，它能与$O_2^-$(超氧阴离子)结合形成毒性更强的超氧亚硝酸($OONO^-$)，超氧亚硝酸质子化后进一步转化成高度反应性的自由基·OH和·NO_2，自由基·OH可用来杀死被吞噬进来的细菌；血管内皮细胞合成的NO，或者与平滑肌接头的神经细胞合成的NO，很容易扩散到在临近的平滑肌细胞中，并激活一种可溶性的GC(与其血红素辅基结合，诱导酶的构象发生变化，提高酶活性达50倍之多)，产生cGMP，cGMP再激活PKA的活性，PKA最终可以导致平滑肌的松弛。

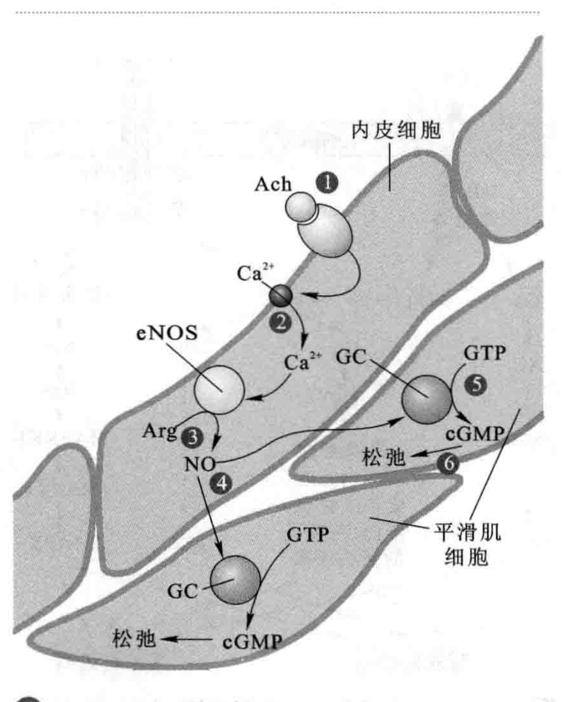


图17–25 NO系统图解

硝化甘油可用来治疗心脏病，原因就是它在体内可转变为NO。NO是一种强劲的血管扩张剂，它在合成以后，通过自由扩散经间隙连接，从内皮细胞进入平滑肌。在这里，NO为平滑肌的第一信使，进入平滑肌细胞后激活GC，产生cGMP。cGMP作为第二信使激活内质网膜上的钙泵，并降低细胞质基质自由Ca^{2+}的浓度，而Ca^{2+}的浓度下降使收缩的肌动蛋白丝/肌球蛋白滑动分开，从而使平滑肌松弛，血管扩张(图17–25)。

Quiz10 万艾可的一种副作用与视觉有关，你认为其中的原因是什么？

在阴茎海绵体中，性刺激可促进 NO 的释放，NO 使 cGMP 产生增加，后者使血液流入阴茎增多而产生勃起。万艾可（Viagra）是现在一种治疗男性性功能障碍即勃起功能障碍的药物，其作用机理是：它在体内可特异性抑制 PDE5 的活性，而 PDE5 主要存在于阴茎海绵体组织中专门水解 cGMP，因此万艾可的存在能够提高 cGMP 在这种组织中的半衰期。

NO 在植物体内也可作为信号分子，但对于植物来说，它还可以通过硝酸还原酶（nitrate reductase）产生。NO 在植物体内的生理作用主要包括：①作为一种发育调节物，能够促进发芽、叶子伸展和根生长，推迟叶子衰老和果实成熟；②作为植物抵抗不容性病原体（incompatible pathogen）的关键信号，诱发抗性相关的高敏感细胞死亡；③能激活一些防御基因的表达，从而在植物获得系统性抗性中起作用；④与蛋白质分子上的半胱氨酸巯基共价结合，使巯基 -SH 转化为 -SNO，从而导致蛋白发生亚硝基化（nitrosylation），这种修饰对多种蛋白质的功能具有调节作用，尤其是对那些参与植物抗病的蛋白质。

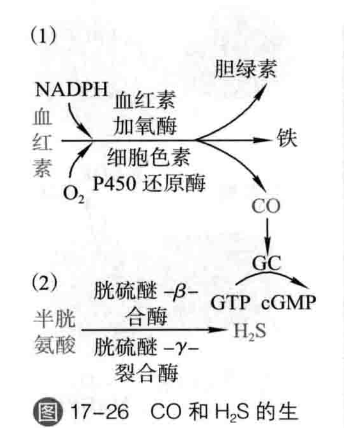


图 17-26 CO 和 H_2S 的生物合成

框 17-2 生化趣事——毒气的妙用

CO 和 H_2S 都曾被一直认为是两种对机体有百害无一益的气体，但现在都被确认为两种重要的气体信号分子，在机体内都可以在特定的酶催化下产生，很容易通过自由扩散的原理到达靶细胞，主要是在循环和神经系统中起作用。

对于 CO 而言，它的重要生理功能包括：调节血管形成、平滑肌细胞分裂、神经转导和血小板凝聚以及抗炎症和抗凋亡。它在机体内的合成与血红素的分解代谢有关（图 17-26），主要由内质网上的血红素加氧酶（heme oxygenase）催化。这种酶存在于人体所有的组织，故体内几乎所有的细胞都能产生 CO。

但与 NO 相比，CO 是相对惰性的气体，但也可以结合并激活细胞内可溶性的 GC，只不过激活的效果要比 NO 差得多。此外，CO 还可以激活与细胞分裂有关的蛋白质激酶 MAPK。实验证明，CO 所具有的所有生理功能都与它能激活 GC 和 MAPK 有关。

就 H_2S 而言，它在体内参与血管扩张、血流调节、白细胞附着、抗炎症和抗凋亡等一系列生理过程。H_2S 主要是由红细胞和血管平滑肌细胞产生，在体内合成的前体主要是半胱氨酸，涉及的酶为胱硫醚 -β- 合酶（cystathionine-β-synthase，CSS）和胱硫醚 -γ- 裂合酶（cystathionine-γ-lyase，CSE）。基因敲除实验表明，缺乏 CSE 的小鼠体内的 H_2S 水平会下降，内皮介导的血管扩张能力减弱，而且表现为高血压。这就证实了 H_2S 对机体的重要性，但究竟是如何起作用的还很不清楚。

以上介绍的几种信号转导通路都十分重要，但它们并非机体内信号转导的全部，事实上，在机体内还有其他很多信号通路，例如 JAK-STAT 信号转导系统、受体离子通道信号转导系统和通过死亡结构域受体起作用的信号转导系统等。这些信号通路受到特定的激素作用下，同样可以被激活，有的产生类似于 RTK 系统的效应，有的导致膜电位的改变，有的则导致细胞凋亡。

e17-4 其他几种信号转导系统

三、信号转导的整合

以上分别介绍了几类重要的细胞跨膜信号转导系统，然而这些系统并不是孤立存在的，而是在它们之间存在着多种形式的相互作用（图 17-27）。

就某一个细胞来说，它在某一时刻所接受的信号刺激绝不可能只有一种，但在不同信号刺激

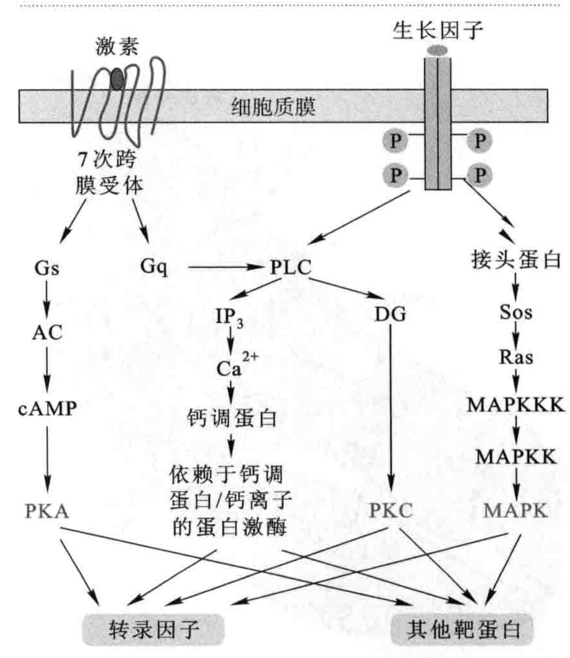
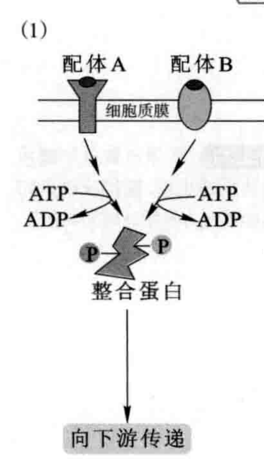


图 17-27 四种平行的信号跨膜转导系统和相互间的联系

下的细胞最终表现出来的却是一种整合效应。细胞在接受不同来源的信号以后，不仅需要进行转导，还需要进行必要的加工和整合，以确定最后的效应是什么，例如细胞是凋亡还是生存？是改变形态还是分泌物质？是分裂还是分化？那么，细胞如何对不同的信号进行整合呢？考虑到前面已阐述的几类主要的信号通路最后都涉及特定的蛋白质激酶，因此激酶在信号的整合中可能起着很重要的作用。

如图 17–28 所示，一个细胞可能存在两种整合机制：一种是两条通路上的蛋白质激酶被激活以后，以同一种蛋白质为底物，但磷酸化位点不同。这种蛋白质在受到来自两条通路上的不同的蛋白质激酶催化以后，即充当整合蛋白，将两条通路上的信号合并整合以后向下游进行传递。Rsk1 就是一种整合蛋白，它可以整合来自 RTK 系统和 PIP_3– 刺激的 PDK1 激酶系统的信号。Rsk1 本身是一种丝氨酸 / 苏氨酸蛋白质激酶，具有两个激酶结构域，两个结构域之间是一段铰链区。RTK 系统中的 ErK 催化 Rsk1 的 C 端激酶结构域和铰链区发生磷酸化修饰，而 PDK1 系统中的 PDK1 催化 Rsk1 的 N 端激酶结构域发生磷酸化修饰。Rsk1 只有在 3 个磷酸化位点都发生磷酸化以后才会被激活。

另一种是两条通路上的蛋白质激酶被激活以后，作用两种不同的蛋白质，而两种不同的蛋白质在分别磷酸化以后，可相互聚合形成二聚体。这样，两条通路上的信号交汇到一种二聚体蛋白质上，通过它的整合，信号再继续向下游传递。



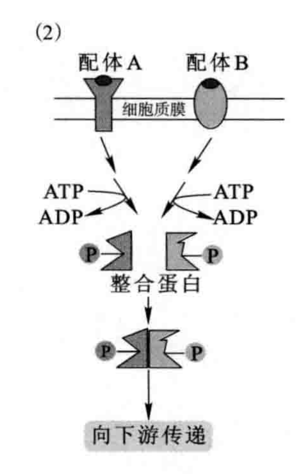


图 17–28 两种细胞内信号转导整合机制

第四节 激素的分泌及其调节

激素作为细胞之间的化学信使，广泛存在于多细胞生物，特别是一些复杂的多细胞生物。一般说来，生物越高等涉及的激素种类就越多。

已在高等动物体内发现近百种以上的激素，它们主要由体内特定的内分泌腺合成或分泌，还可能由分散在全身各处的一些特定的组织细胞分泌。而在植物体内也发现了至少 8 种类型的激素。与动物激素相比，植物激素仅仅是一些刺激生长或抑制生长的小分子物质，而不是内环境稳定的调节物。所有植物细胞都能产生植物激素，不像动物那样通常由专门的特异性器官分泌。此外，植物激素一般不作用于较远的组织或细胞，而是作用于临近的组织和细胞。

e17–5 动物激素与植物激素

动物激素的生理效应在很大程度上取决于它们在血液和细胞外液中的浓度。任何一种激素的浓度过高或过低几乎不可避免地将导致疾病的发生，各种内分泌失调引起的病例不胜枚举，因此机体对激素分泌的精确控制就显得十分重要。

作用于靶细胞的激素有效浓度主要由三种因素决定：①产量，由合成和分泌决定，这是控制激素浓度最重要的环节。②转运速率，显然转运速率越快，分泌的激素越容易达到靶细胞。③半衰期，激素与其他生物分子一样，在体内会被降解或被排泄，因此有一定的半衰期。关闭激素的分泌只会导致半衰期短的激素循环浓度迅速下降。如果一种激素的半衰期较长，即使关闭它的分泌，其有效浓度仍然会持续一段时期。

以高等动物激素为例：其释放在原则上不是持续不断的，而是短促暴发式（short burst）的，在前后两次暴发式释放之间，很少有激素分泌。主要有 3 类刺激可导致激素的暴发式释放：①来自神经系统的动作电位；②血浆和细胞外液的化学变化；③其他激素。因此，控制激素分泌的手段首先是来自神经系统的控制，其次是由于血浆和细胞外液的化学变化引发的反馈（主要是负反馈），最后就是受其他激素的控制。这 3 种手段在动物体的下丘脑 – 垂体 – 靶腺组成的控制轴上被有机地结合在一起。

图 17–29 反映了 3 种调控方式之间的关系，下丘脑处于调控网络的最上端，它受中枢神经系统（central nervous system，CNS）的控制（通过边缘系统），所分泌的释放激素或释放抑制激素控制脑垂体激素的分泌，而脑垂体分泌的激素又控制其他内分泌腺即靶腺的激素分泌。当靶腺分泌的激素浓度达到一定水平以后，负反馈环路开始起作用。共有 3 种负反馈通路：①长反馈（long feedback loop），是

指靶腺分泌的最终激素与CNS、下丘脑或脑垂体上内源的受体结合，阻止激素从这些细胞的释放；②短反馈（short feedback loop），是指脑垂体分泌的激素与下丘脑上的内源受体结合，最终抑制下丘脑分泌释放激素；③超短反馈（ultra-short feedback loop），是指下丘脑分泌的释放激素抑制自身的分泌。

Quiz11 利用激素分泌调控的负反馈机制，解释缺碘是如何导致地方性甲状腺肿大的？

以性激素为例，其分泌受脑垂体分泌的促性腺激素（FSH和LH）的促进，而促性腺激素的分泌又受到下丘脑分泌的GnRH的促进。血浆中的性激素在浓度达到一定阈值以后，便以负反馈的方式抑制脑垂体分泌促性腺激素和下丘脑分泌GnRH，最终导致性激素分泌的抑制。而随着性激素的代谢或排泄，浓度开始下降。在下降到阈值以下后，负反馈解除，下丘脑又重新开始分泌GnRH，GnRH再次导致促性腺激素的释放，促性腺激素又导致性激素的释放。

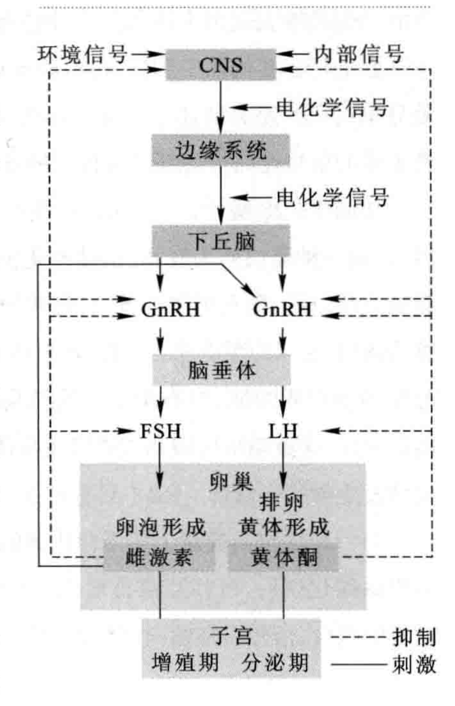


图17-29 位于下丘脑-垂体轴上激素分泌的调节

其他一些激素，如胰岛素，虽然不在下丘脑-垂体轴上，但仍然会受到负反馈的控制。例如，当喝一杯牛奶或吃一块糖果以后，体内将会发生一系列的事件：摄入的乳糖或蔗糖被水解以后，释放出的葡萄糖被小肠吸收进入血流，导致血糖浓度提高；血糖水平升高刺激胰岛β细胞分泌胰岛素；胰岛素促进葡萄糖转运到许多体细胞，血糖浓度开始下降；血糖浓度下降到一定阈值，促进胰岛素分泌的刺激消失，这时抑制胰岛素分泌的负反馈机制开始起作用，胰岛素也就不再分泌。

e17-6 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Hormone（维基百科有关激素的内容）
2. http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/S/SexHormones.html（免费的生物学网站有关性激素的内容）
3. http://themedicalbiochemistrypage.org/steroid-hormones.html（完全免费的医学生物化学课程网站有关固醇类激素的内容）
4. http://themedicalbiochemistrypage.org/peptide-hormones.html（完全免费的医学生物化学课程网站有关肽类激素的内容）

参考文献：

1. Ryba N J, et al. The receptors and coding logic for bitter taste. *Nature*, 434: 225–229, 2005.
2. Luo H, Charles R. The JAK/STAT pathway and *Drosophila* development. *BioEssays*, 23: 1 138–1 147, 2001.
3. Buck L B. The molecular architecture of odor and pheromone sensing in mammals. *Cell*, 100(6): 611–618, 2000.
4. Malnic B, et al. Combinatorial receptor codes for odors. *Cell*, 96(5): 713–23, 1999.
5. Bredt D S, Snyder S H. Nitric oxide: A physiologic messenger molecule. *Annu Rev Biochem*, 63: 175–195, 1994.
6. Strader C D, et al. Structure and function of G protein-coupled receptors. *Annu Rev Biochem*, 63: 101–132, 1994.

结构生物化学思考题

1. 多聚 Leu 在疏水环境中比在水中更容易形成 α 螺旋

(1) 为什么疏水环境有利于多聚 Leu 形成螺旋?

(2) Ile 的侧链与 Leu 含有数目完全相同的 C 原子,然而多聚 Ile 在疏水环境中不能形成稳定的螺旋,为什么?

2. 正协同效应、别构效应和波尔效应对 Hb 的生理功能有何影响? 为什么 Mb 没有上述三种效应?

3. 鳄鱼能够潜伏在水下而不呼吸达 1 h 之久,捕猎时会将呼吸空气的大型猎物拖入水中溺死。与此生活方式相适应的是,它几乎能够利用血液中 100% 的氧气,而人只能抽取血液中约 65% 的氧气。鳄鱼的血红蛋白不结合 2,3-BPG,但其脱氧血红蛋白更喜欢结合 HCO_3^-。试解释这为什么有助于鳄鱼获取食物?

4. HbF 与 HbA 在结构、功能以及氧合曲线上有何差别? 镰状细胞贫血不会影响还没出生的婴儿,为什么? 用你学过的知识,想出一种合理的治疗镰状细胞贫血的方法。

5. 有人对从血库里贮存较久的血液中,抽取了血红蛋白,并对其氧合曲线进行了专门的研究,发现氧合曲线不再是 S 形,为什么? 你认为这样的血液还能否使用?

6. 细胞在受热的情况下,会受到热激蛋白(HSP)的保护。HSP 通过与在温度突然提高时暴露出来的氨基酸残基侧链基团的结合保护细胞。

(1) 列出可能与 HSP 结合的氨基酸残基。

(2) 这些氨基酸具有的什么性质使得它们能够与 HSP 结合?

(3) HSP 与它们结合为什么能够保护细胞?

7. 已发现许多蛋白质在室温下稳定,但在较低温度下却不稳定,从而导致所谓的冷变性。你认为导致冷变性的原因是什么?

8. 假定你从人脑中分离纯化到一种新的蛋白质,你如何测定出它的大小和 p*I*? 如何知道它在肝细胞里是否表达? 如何测定它的三级结构? 如何预测出它可能具有的生物学功能?

9. 如果你是一名内科医生,遇到一个急性胰腺炎患者,其体内由胰腺分泌的蛋白酶在细胞内提前激活,你会如何治疗这个患者? 是让他服用胰蛋白酶抑制剂或胰凝乳蛋白酶抑制剂,还是弹性蛋白酶抑制剂? 为什么?

10. 为什么胰蛋白酶、胰凝乳蛋白酶和弹性蛋白酶的特异性差别很大,但催化机制完全相同? 如果它们被分泌到胃液之中,而不是肠液中,它们还会有正常的活性吗? 为什么?

11. 如何区分辅基和辅酶? 有哪些辅酶或辅基分子中含有腺苷? 这些含有腺苷辅酶的存在对于你理解生命的起源有何启示? 为什么?

12. 参与细菌 tRNA 前体 5′端后加工的核糖核酸酶 P 已被证明是一种核酶,它由蛋白质和 RNA 两种成分组成。在比较各种不同细菌的核糖核酸酶 P 的一级结构和高级结构后发现,U69 是绝对保守的,这个位置的 U 不与任何碱基配对,而是以突起的形式游离在 RNA 二级结构之中,为了研究它与酶活性的关系,有人对 U69 进行了定点突变,并对突变体的速度常数 k 和 K_m 进行了测定,结果如下表:

M1 RNA	k/min^{-1}	$K_m/(\text{nmol}\cdot\text{L}^{-1})$
野生型	0.26	1.7
U69→A69	0.062	4
U69→G69	0.003 4	73
U69→C69	0.005 6	3
U69 缺失	0.005 6	7
U69+U69	0.005 4	181

根据表中的结果，你认为U69对于酶与底物的结合还是对于酶的催化更为重要？如果在U69后面再引入一个U(U70)，以扩大突起的面，你认为结果又是如何？

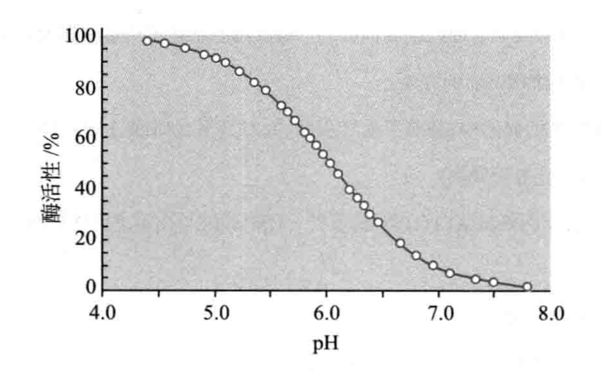
13. 有一种酶E作用底物S，产生产物P。化合物C是此酶的别构激活剂。使用定点突变技术，可将酶E的Val57突变成其他几种氨基酸残基。野生型和变体被纯化以后，分别在有C和无C的条件下测活，活性分析的结果见下表：

酶E	无C时的酶活性	有C时的酶活性
野生型	10.3	51.4
Val57→Ser57	10.5	30.2
Val57→Glu57	10.2	11.1
Val57→Ala57	10.1	49.5

(1) 根据突变体的数据，你认为Val57位于别构中心，还是活性中心？

(2) 根据氨基酸的性质，解释突变如何影响到酶E的活性。

14. 一种酶的活性受到pH的影响，其与pH的关系如下图：



(1) 你认为最有可能是哪一种氨基酸侧链基团的表观pK_a影响酶的活性？

(2) 该基团的质子化还是去质子化形式才是酶活性所需要的？为什么？

15. 有两种微生物，一种生活在热泉中，另外一种生活在南极，你认为这两种微生物的基因组DNA在一级结构和三级结构上会有何种差别？为什么？

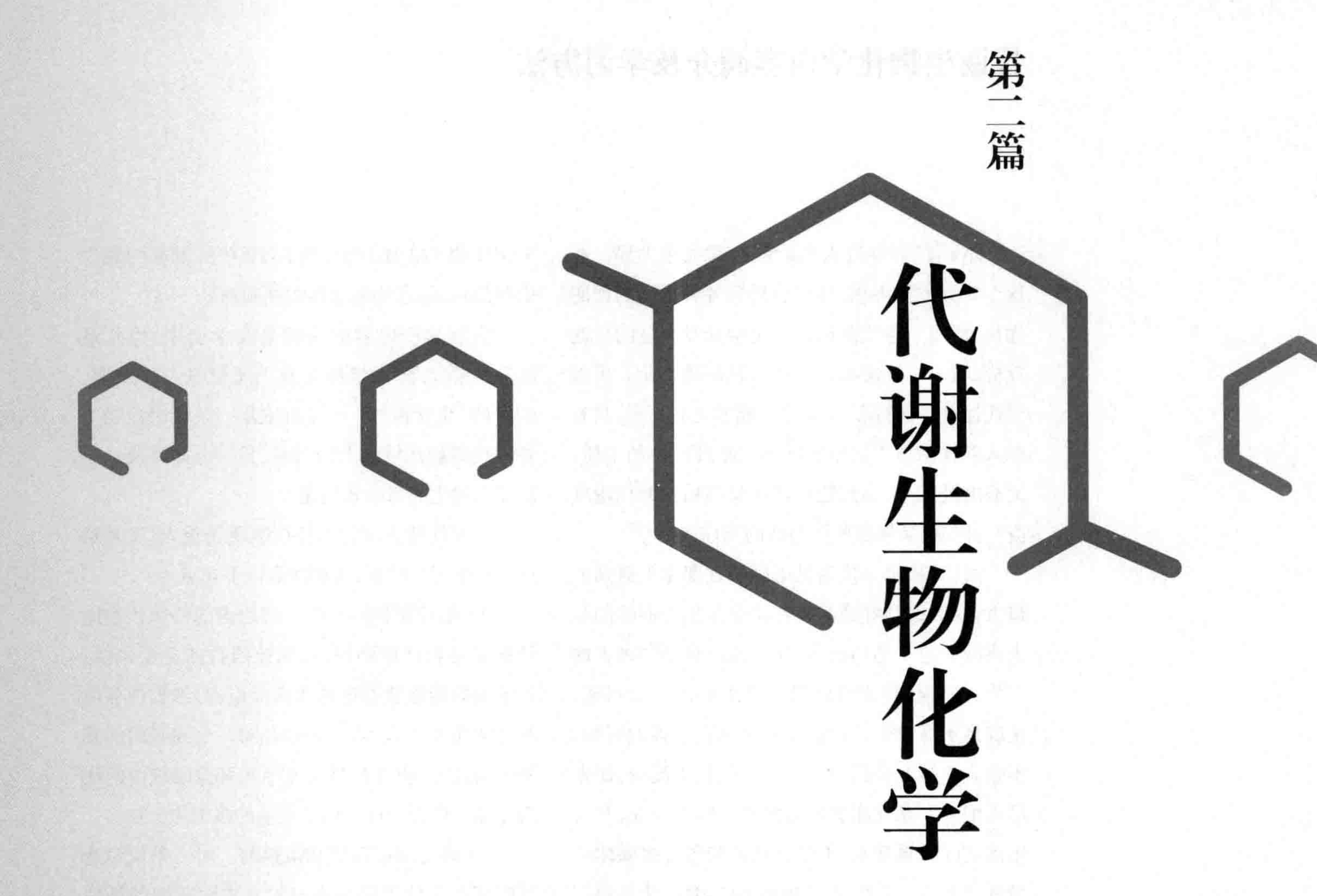
16. DNA双螺旋在什么情况下会形成超螺旋？若将共价闭环负超螺旋和正超螺旋DNA分别与核酸酶S1保温在一起，这对两种超螺旋DNA的结构有没有影响？为什么？

17. 不能自身合成脂肪酸的细菌突变体，需要在培养基之中提供外源的脂肪酸。假定有两种细菌培养物，每一种培养基之中都含有几种直链脂肪酸的混合物，既有饱和的也有不饱和的。脂肪酸的碳链长度为10~20。如果一种培养物的培养温度是18℃，另一种是40℃，那么经过几代培养以后，两种培养物中的细菌在膜脂组成上会有何不同？

18. 葡糖氧化酶催化葡萄糖的氧化。此酶对于β-D-葡萄糖是高度特异性的，对α-D-葡萄糖不起反应。然而，这种酶经常被临床上用来测定血糖(β-D-葡萄糖和α-D-葡萄糖的混合物)浓度，为什么？

19. 有两种激素(激素A和激素B)，激素A作用于靶细胞以后并不引发胞内cAMP浓度的变化，激素B作用于相同的靶细胞以后，可导致胞内的cAMP水平的上升，但上升的幅度并不很高。然而，如果将激素A与激素B同时作用于靶细胞，则可导致cAMP水平的急剧升高。为什么？

20. Ras蛋白的突变可导致一系列癌症的发生。其中最常见的一种突变是1个Gly突变成Val。为什么这种突变可导致癌症？



第二篇

代谢生物化学

代谢生物化学内容简介及学习方法

本篇将带你进入色彩缤纷的代谢世界。在这个奇妙的世界里，你将会遇到各种不同的代谢途径，而每一条代谢途径一定会涉及大量的代谢反应，每一个反应都由一种特定的酶催化。可以说代谢世界既精彩，又无奈。精彩之处在于，只有深入到代谢世界之中才能真正把握生命的本质；无奈之处在于，如何能够记住那些枯燥的代谢反应？又如何掌握那些限速酶的调控机制？

可以说，讲好代谢是老师们在课堂上遇到的最大挑战，而学好这些内容又是学生们面临的最大挑战。对于老师而言，在讲授这些章节内容的时候，一定要想方设法激发学生们的学习兴趣，正确地引导学生。这就需要老师在讲课的时候，不是去单纯地介绍一个又一个的代谢反应，而是尽可能将一条代谢途径有机地串联在一起，并与生活实际联系起来，再结合动画和各种多媒体手段进行讲解。例如，在讲糖酵解的时候，应该将它与乳酸发酵和酒精发酵联系在一起，可组织学生在暑假自酿葡萄酒。在讲 TCA 循环的时候，可以将它与武大郎被砒霜毒死的故事联系在一起。在讲脂代谢的时候，注意将它与肥胖联系在一起。在讲氧化磷酸化的时候，要用动画表现呼吸链传递电子和创造质子梯度的过程，用动画和照片及录像显示 F_1F_0-ATP 合酶的结合变化机制。而对于学生来说，第一要树立信心，千万不要被大量的生化反应吓到！第二，要区分哪些内容是必须掌握的，哪些内容只要略知一二，哪些内容根本就不需要记。第三就是注意学习方法，这一点是最重要的。

在学习一条代谢途径的时候，需要掌握以下内容：①代谢途径的场所和功能；②进入和离开代谢途径的各种代谢物的名称；③一条代谢途径如何跟其他代谢途径进行整合和协调；④代谢途径的限速步骤、限速酶的名称和调节方式；⑤调节限速酶活性的各种别构效应物的名称，能够判断是作为正别构效应物还是负别构效应物起作用；⑥生成或消耗 ATP 或者等价物（如 GTP、NADH 和 $FADH_2$）的反应；⑦某些关键酶的缺失引起的后果；⑧代谢途径的抑制剂。

掌握这些内容的关键是要学会用“生化逻辑”去理解，而不是孤立地去死记每一步反应。这里的“生化逻辑”一方面涉及一条代谢途径上各种代谢物在结构上的关联，另一方面涉及每一步反应发生的目的和用途。

一条代谢途径的调控机制既是重点，又是难点。在学习的时候，需要注意以下几点：

（1）明确调控的目的　细胞内的任何代谢途径都是受到严格调控的，调控的目的主要包括：对细胞的能量状态迅速作出反应；对细胞内环境的变化作出反应；协调和整合同一个细胞内的其他代谢途径；协调和整合发生在其他细胞内的代谢途径。最后一点适用于复杂的多细胞生物。

（2）确定限速反应和限速酶　对一条代谢途径调节的主要手段是通过对其中限速酶的活性进行全方位的控制。因此，首先需要确定一条代谢途径的限速步骤。那么，什么样的反应才构成限速步骤呢？一般是不可逆反应，因为不可逆反应才是这条代谢途径所特有的，控制它才不会波及到其他代谢途径。那么什么样的反应是不可逆反应呢？这主要与反应自由能的变化有关系。如果一个反应的自由能变化是一个很大的负值，这往往就意味着该反应很可能是不可逆的。有时，一条代谢途径有几步不可逆反应，这时机体会选择其中一步作为重点调控的对象。例如，糖酵解有三步不可逆反应，但只选择第二步不可逆反应作为最重要的调控位点。那么，如何判断细胞会选择哪一步作为最重要的调控步骤呢？这里有两个规则：一是看哪一步反应是这一条代谢途径所特有的？二是这一步反应所处的位置。一般而言，细胞会选择代谢途径特有的、靠前的不可逆反应作为重点调控的对象。

（3）确定调节的方式　对于限速酶活性的调节方式主要是酶的量变和质变，这在酶学一章已有详细的介绍，不过不同的限速酶选用的方式不

一定相同。

(4) 如果是别构调节,细胞会选择哪些代谢物作为别构效应物? 别构调节是细胞调节限速酶活性的一种最常见的手段,所选用的别构效应物一般包括:细胞能量状态的指示剂,如 ATP、AMP 和 ADP。以它们作为别构效应物的酶一般参与细胞内的产能代谢途径或储能代谢途径;一条代谢途径的终产物,以反馈抑制的方式起作用;一条代谢途径的起始底物,以前馈激活的形式起作用;其他代谢途径的中间物,以协调功能相关的代谢途径。例如,TCA 循环产生的柠檬酸,可作为脂肪酸合成限速酶和糖酵解限速酶的别构效应物,脂肪酸氧化产生的乙酰 -CoA,可作为糖异生限速酶的别构效应物;限速酶本身的产物,这是一种短反馈。

(5) 如果是共价修饰,哪一种形式的酶有活性? 共价修饰的调节方式主要是可逆的蛋白质磷酸化,如何能够正确判断一种酶磷酸化以后被激活还是被抑制呢? 为此我总结出了一条规则,只要记住两句话就行了:一是饥饿的情况下我们体内的蛋白质磷酸化水平提高(饥饿的情况下肾上腺素和胰高血糖素分泌提高,导致蛋白质的磷酸化);二是血糖浓度的稳定压倒一切(我们的脑细胞严重依赖于血糖提供能量)。下面以糖原磷酸化酶和糖原合酶为例加以说明:假定你现在在上上午第四节课了,你饥饿难忍,于是你体内的蛋白质磷酸化水平提高,这包括糖原磷酸化酶和糖原合酶。由于饥饿,体内血糖的浓度在下降,为了稳定血糖浓度,显然需要抑制糖原的合成,激活糖原的分解,这样可以让释放出来的葡萄糖单位进入血液,稳定血糖。而催化糖原合成的酶是糖原合酶,分解糖原的酶是糖原磷酸化酶。既然在饥饿的情况下,这两种酶都被磷酸化了,显然磷酸化的糖原磷酸化酶被激活,磷酸化的糖原合酶被抑制才有利于血糖浓度的稳定。

(6) 注意原核生物和真核生物的差别 同样的代谢途径在原核细胞和真核细胞内的调控机制会有很大差别,因此需要区别对待。但在教材上介绍的多为发生在高等动物体内的调控机制。

第十八章　代谢总论

代谢(metabolism)是生命最基本的特征之一,它是指生物体内发生的所有化学反应,包括物质代谢和能量代谢两个方面的内容。由于在物质代谢的同时总伴随着能量的变化,因此物质代谢和能量代谢是不可分割的。

本章将重点介绍代谢的基本概念、代谢的基本特征、代谢研究的主要方法和代谢组学。

第一节　代谢的基本概念与特征

一、代谢的基本概念

细胞内发生的代谢反应几乎都是在酶的催化下完成的,一种物质在细胞内的转变,无论是分解还是合成,通常由一系列酶促反应构成,各个反应按照一定的次序有条不紊地进行,构成特定的代谢途径(metabolic pathway)。在一个细胞里,有多条代谢途径。这些代谢途径既相互独立,又相互关联,形成一种复杂精妙的代谢网络(图 18-1)。在一条代谢途径之中,前一个酶的产物刚好作为后一个酶的底物,很难孤立地把它们归为底物还是产物,一般就称其为代谢物(metabolite)或代谢中间物(metabolic

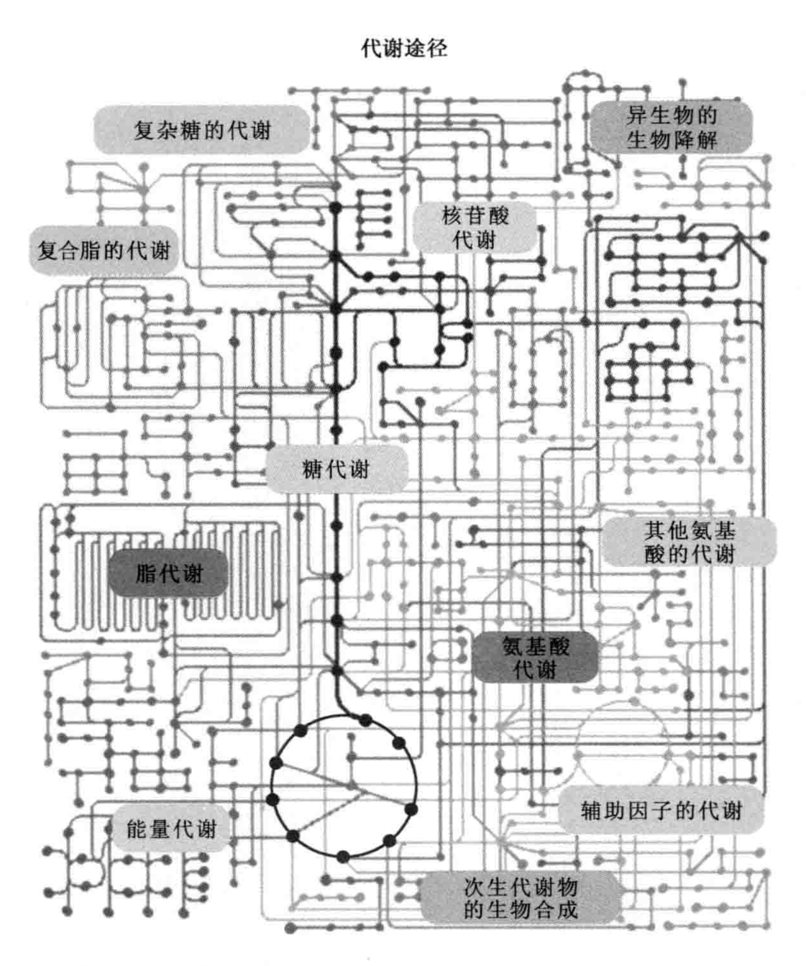


图 18-1　细胞内的代谢途径和代谢网络

intermediate)。

一条代谢途径中的酶可以通过三种方式组织在一起(图 18-2):

(1) 分散形式　这种方式的效率最低,因为前一个酶释放的产物在与后一个酶结合之前已被环境介质所稀释。

(2) 多酶复合体或酶系　这种方式较为普遍,由于在反应之中,前一个酶释放的产物可以直接与后一个酶的活性中心结合,因此反应效率得以大幅度提高。此外,这种方式还有利于调控。比较典型的例子有丙酮酸脱氢酶系、α- 酮戊二酸脱氢酶系和原核细胞的脂肪酸合成酶系等。

(3) 与膜结合的多酶复合物(membrane-bound multi-enzyme complex)　参与这一类代谢途径的酶按照一定的方向和次序整合在膜上,以膜为平台,例如生物氧化中的呼吸链和光合作用中的光合链。

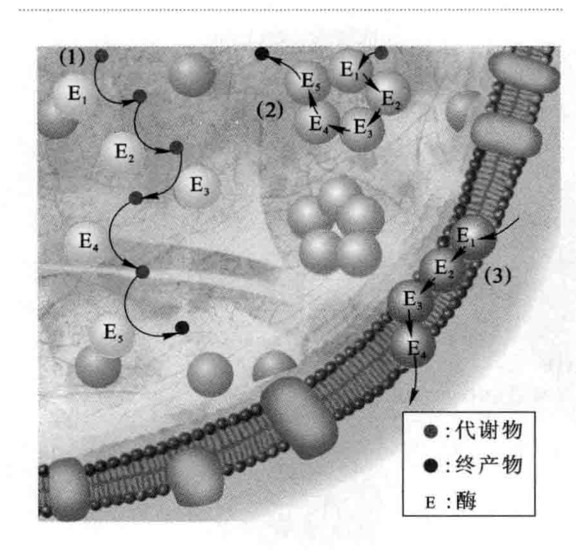


图 18-2　酶的三种组织方式

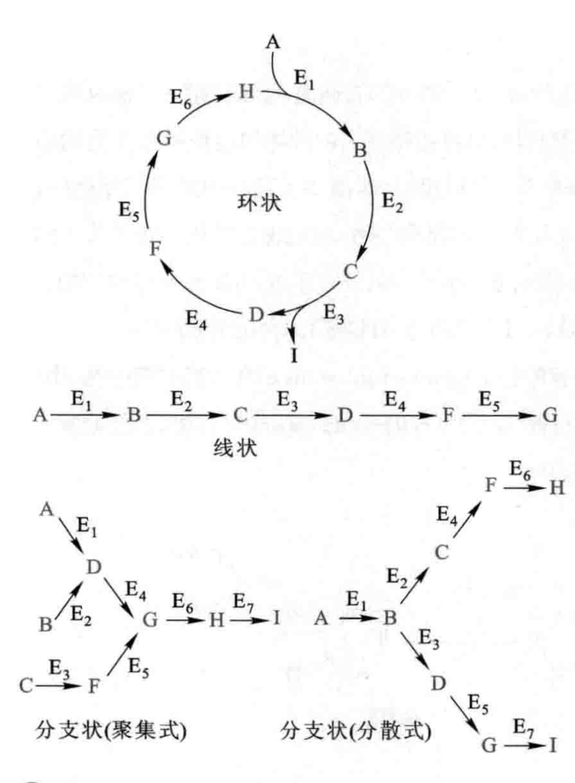


图 18-3　代谢的三种途径

有时,某一条代谢途径上的某几个酶活性、甚至所有酶活性融合在一条多肽链上或者由同一个蛋白质的不同亚基承担。例如,哺乳动物的脂肪酸合酶由两条相同的肽链组成,每条肽链竟然具有 7 个酶活性!酶以这种方式组织在一起,不仅大大提高了催化的效率,有利于对各个酶活性的调控,还增加了基因的编码能力,也方便了对这些酶基因表达的调控。

按照代谢进行的方向,代谢途径可以分为(图 18-3):线状(linear),如糖酵解;环状(cyclic),如三羧酸循环、卡尔文循环和尿素循环;分支状(branched),如许多氨基酸的合成(参看第三十一章"氨基酸代谢")。只有在循环代谢途径中,才涉及一种起始代谢物在最后一步反应得以再生。例如三羧酸循环中的草酰乙酸,卡尔文循环中的 1,5- 二磷酸核酮糖和尿酸循环中的鸟氨酸。

细胞内的所有代谢反应都可以汇入到 3 种代谢途径,即分解代谢(catabolism)、合成代谢(anabolism)(图 18-4)和不定向代谢(amphibolic metabolism)之中。

分解代谢为复杂的代谢物转变为简单的代谢物的过程,通常由 3 个阶段组成:

第一阶段,复杂的大分子(如淀粉、糖原、蛋白质、核酸和脂肪等)分解成它们的组成单位,如单糖、氨基酸、核苷酸、甘油和脂肪酸等。在此阶段,一般没有 ATP 的合成。

第二阶段,第一阶段形成的小分子组成单位,进一步降解成更小的分子,如乙酰辅酶 A 等。在此阶段一般有少量的 ATP 合成。

第三阶段,由最终的共同代谢途径组成,包括三羧酸循环、呼吸链和氧化磷酸化。经过此阶段的反应,代谢物被彻底氧化为 H_2O 和 CO_2,同时生成大量的 ATP 或它的等价物。

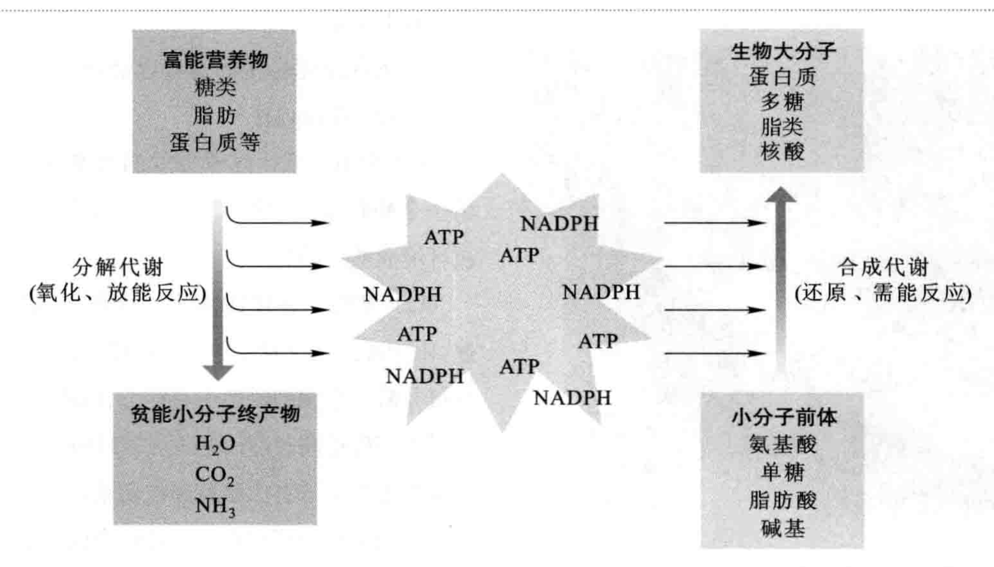


图 18-4 分解代谢和合成代谢

框 18-1 生化聚焦——底物通道运输

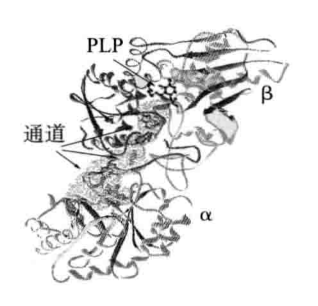


图 18-6 大肠杆菌色氨酸合酶两个活性中心直接的底物通道

在某些多酶复合体和多功能酶的各个活性中心之间，存在一种专门的通道，以便让前一个酶反应产生的中间代谢物直接转移到下一个酶的活性中心并被利用，这种运输反应中间物的过程称为底物通道运输（substrate channeling）。底物通道运输有许多好处，它们包括：①减少反应中间物的运输时间；②由于中间物没有离开酶而释放到溶液中，因此可防止反应中间物因扩散作用造成的损失；③使不稳定的中间物接触不到溶剂；④隔离对细胞有毒性的反应中间物；⑤克服不利的反应平衡；⑥阻止反应中间物流入到其他竞争性代谢途径；⑦防止过量反应中间物的堆积；⑧有利于协调对整条代谢途径的调节。

底物通道运输机制的典型例子有：大肠杆菌的色氨酸合酶（tryptophan synthase）和大肠杆菌的氨甲酰磷酸合成酶（carbamoyl-phosphate synthetase，CPS），以及利什曼原虫（*Leishmania major*）的二氢叶酸还原酶－胸苷酸合酶（dihydrofolate reductase-thymidylate synthase）。

吲哚 -3- 磷酸甘油 —(α；3- 磷酸甘油醛)→ 吲哚 —(β；丝氨酸)→ 色氨酸

图 18-5 大肠杆菌色氨酸合酶两个活性中心催化的反应

以大肠杆菌的色氨酸合酶为例，它由 2 个 α 亚基和 2 个 β 亚基组成，α 亚基和 β 亚基上各有一个活性中心，但催化不同的反应（图 18-5）。其中 α 亚基催化吲哚 -3- 磷酸甘油（indole-3-glycerol phosphate，IGP）形成吲哚和 3- 磷酸甘油醛。β 亚基催化吲哚和丝氨酸形成色氨酸，这一步反应需要磷酸吡哆醛（PLP）。两个亚基的活性中心通过 2.5 nm 长的疏水通道相连，第一步反应发生在 α 亚基的活性中心，有两个产物——吲哚和 3- 磷酸甘油醛。其中吲哚直接通过这个疏水的通道进入 β 亚基的活性中心（图 18-6），随后作为底物之一参与第二步反应，而 3- 磷酸甘油醛为高度亲水的分子，无法通过疏水的通道，于是被直接释放出去。

合成代谢由一系列的生物合成反应构成，在消耗能量或输入能量的前提下，通过合成代谢，简单的分子变成了复杂的分子，小分子单位变成了大分子物质。能量的输入有两种方式：一是分解代谢中产生的 ATP，二是以 NADPH 形式存在的高能电子。

不定向代谢是指细胞内某些具有双重功能的代谢途径，它既可以用于分解代谢，又可以用于合成代谢，例如三羧酸循环。

Quiz1 一条分解代谢途径和一条合成代谢途径的总自由能变化都是负值吗？为什么？

二、代谢的基本特征

生物体内的代谢反应具有以下特征：

(1) 反应条件温和　生物体内发生的代谢反应通常在较温和的条件下进行（37℃，101.3 kPa，pH 7），这是因为有酶的催化。

(2) 高度调控　细胞内的代谢途径并不是以恒定不变的速率进行，有时会加快，有时则受到抑制，这是因为细胞内存在复杂多变的调控机制（详见第十二章“酶活性的调节”以及其他章节中与代谢调控相关的内容）。

(3) 每一个代谢途径都是不可逆的　尽管每一条代谢途径之中的多数反应为可逆反应，可是总会有一两步反应由于 ΔG 为极大的负值而不可逆。例如，糖酵解总共有 10 步反应，有 3 步是不可逆反应。不可逆反应的存在决定了每一条代谢途径的单向性，这更有利于机体对代谢进行调节。

Quiz2 细胞内为什么会有不可逆反应？这是由于催化反应的酶只催化正反应，不催化逆反应吗？

(4) 一个代谢途径至少存在 1 个限速步骤　代谢的调节并不需要对一条途径上所有的酶促反应进行控制，只需要对其中的一两步关键步骤或限速（rate-limiting）反应进行调控即可。限速反应一般是代谢途径的第一步不可逆反应。

(5) 各种生物在基本的代谢途径上是高度保守的　不论是哪一种生物，其基本的代谢途径都是极为相似的，例如，糖酵解途径存在于所有的细胞之中，而且反应的性质非常接近。

代谢途径的高度保守性不仅能够说明所有的生物在进化上有一个共同的祖先，还为人们研究各种代谢途径提供了方便，我们完全可以从一个简单的生物系统着手去研究一条代谢途径或其中的某一步酶促反应。

(6) 代谢途径在真核细胞中是高度分室化的　真核细胞其内部结构是高度分室化的（highly compartmentalized）。分室化将不同的代谢途径限定在不同的区域，这不但有利于调控，而且能够防止一个反应在错误的时间发生在错误的地点（例如细胞内的水解酶如果不集中在溶酶体，细胞就会自溶而死），还能减少细胞内发生“无效循环”（futile cycle）的可能性。下面就“无效循环”这种情况举例说明（图 18-7），细胞内存在两个方向相反的不可逆反应，第一个反应发生在细胞质基质，是糖酵解途径中的第一步反应，此反应在消耗 ATP 的情况下，葡萄糖被激活为 6- 磷酸葡糖。第二个反应发生在内质网，为糖异生途径中的最后一步反应，反应将 6- 磷酸葡糖水解成葡萄糖。试想一下，如果这两个反应发生在同一地点，后果会是什么？显然，在不断消耗 ATP 的情况下，6- 磷酸葡糖难以形成，细胞会陷入一种“无效循环”的状态，ATP 分子中的高能键白白浪费，变成了热。

$$\text{葡萄糖} + \text{ATP} \xrightarrow[\text{发生在细胞质基质}]{\text{己糖激酶}(Mg^{2+})} \text{6- 磷酸葡糖} + \text{ADP}$$

$$\text{6- 磷酸葡糖} \xrightarrow[\text{发生在内质网}]{\text{6- 磷酸葡糖磷酸酶}} \text{葡萄糖} + \text{Pi}$$

图 18-7　细胞内可能存在的一种“无效循环”

表 18-1 显示出真核细胞内主要代谢途径的分室化情况。必须指出，细胞内某些代谢途径中的部

表 18-1　代谢途径的分室化

代谢途径	发生区域
三羧酸循环、氧化磷酸化，脂肪酸氧化，氨基酸分解	线粒体
糖酵解、脂肪酸合成、磷酸戊糖途径	细胞质基质
DNA 复制、转录、转录后加工	细胞核、线粒体、叶绿体
膜蛋白和分泌蛋白的合成	糙面内质网
脂肪、磷脂和胆固醇的合成	光面内质网
翻译后加工（糖基化）	内质网和高尔基体
尿素循环	肝细胞线粒体和细胞质基质

Quiz3 哺乳动物成熟的红细胞还保留哪些代谢途径?

分反应发生在一个区域,而另外一些反应发生在其他区域。例如,糖异生起始阶段的几步反应发生在线粒体,此后大多数反应发生在细胞质基质,而最后一步反应又发生在内质网。

三、代谢中的氧化还原反应和氧气在代谢中的作用

细胞中的很多代谢反应为氧化还原反应(详见第二十章“生物氧化”),通常认为分解代谢为氧化性的(oxidative),这是因为代谢物在其中被氧化,并以 $NAD(P)^+$ 或 FAD 为氢和电子受体,而合成代谢为还原性的(reducing),这是因为在生物合成反应中,NADPH 作为还原剂提供生物还原能力,形成还原性高的产物。

不同的生物体在代谢反应中对 O_2 的需求不太一样,那些在分解代谢途径之中以 O_2 作为最终电子受体的一类生物通称为需氧生物(aerobe),否则就是厌氧生物(anaerobe)。需氧生物又分为专性需氧生物(obligate aerobe)和兼性需氧生物(facultative aerobe),前者必须使用 O_2 作为最终的电子受体,后者则取决于 O_2 的可得性,即在有氧的环境下以 O_2 为最终电子受体,O_2 缺乏则以其他物质(如丙酮酸)作为最终电子受体。厌氧生物也可以分为专性厌氧生物和兼性厌氧生物,前者必须生活在无氧的环境之中,只能以 O_2 以外的物质作为最终的电子受体。

需氧生物体内的某些代谢途径并没有 O_2 的参与,例如糖酵解。从进化的角度来说,这些无氧代谢途径的出现应该先于有氧代谢途径。

第二节　代谢研究的主要内容和方法

代谢研究的主要内容包括以下几个方面:①确定参与每一个代谢反应的酶与辅酶的结构与功能,这需要对有关的酶进行分离、纯化、定性和定量的研究。②确定一条代谢途径中的底物、中间代谢物和终产物的结构、名称和反应的类型。获取这些信息对于了解一个反应的机制非常重要。③确定一个酶促反应的调节机制。

Quiz4 什么是回体法?举一例加以说明。

在进行代谢研究之前,必须选择好合适的研究对象:是以微生物作为研究对象,还是以动物为研究对象?是在整个生物体,还是以其中的一个器官、组织或在亚细胞水平去研究代谢?是在体内(*in vivo*),还是在体外(*in vitro*)或用回体法(*ex vivo*)进行研究?

在确定好研究对象和目的以后,就要选择合适的方法。代谢途径的阐明与研究方法的进步是分不开的。纵览代谢研究的历史,下面几种方法曾在代谢研究中立下汗马功劳。

(1) 同位素示踪　在代谢研究之中,使用最多的方法要数同位素示踪(isotope tracer)。实际上,细胞内主要的代谢途径都是通过此方法得以阐明的,如糖酵解、三羧酸循环和卡尔文循环等。此方法需要使用同位素标记特定的代谢物,然后将被标记的分子引入细胞,通过追踪它们在生物体内的去向和代谢转变,从而确定与被标记物有关的代谢途径。经常使用的同位素有放射性和非放射性两类,前者包括 ^{14}C、^{32}P、^{35}S 和 ^{3}H 等,后者包括 ^{18}O 和 ^{15}N 等。

Quiz5 若是需要用放射性同位素追踪 DNA 的复制,那你如何对 dNTP 进行放射性同位素标记?

(2) 代谢抑制剂的使用　既然代谢反应离不开酶,那么总可以筛选到一种酶的抑制剂,去阻断一个特定的代谢反应。如果一条代谢途径上的某一步反应被抑制以后,就必然导致这一步反应的底物以及它前面所有的代谢物在细胞内的堆积。综合不同抑制剂(图 18–8 所示的 I_1、I_2、I_3、I_4 和 I_5)的作用结果,可以将一条完整的代谢途径大致地勾画出来。

图 18–8　使用代谢抑制剂确定代谢途径

(3) 代谢遗传缺陷型突变体的使用　使用代谢遗传缺陷(genetic defect)型突变体与使用酶抑制剂的效果是一样的。一个酶基因的缺陷会导致其活性的抑制或丧失,这与使用抑制剂的结果类似,它同样会导致特定代谢物的堆积。

(4) 基因操作　主要是使用转基因技术、基因敲除或敲减技术和基因在某些组织中的定向表达等手段(参看第四十四章"重组 DNA 技术"),来确定某一种酶在细胞之中的功能以及它在代谢中所发挥的作用。随着基因组计划(genome project)和生物信息学的兴起以及高通量基因表达(high throughput gene expression)方法的建立,代谢研究又多了新的技术和方法。

随着基因组学和蛋白质组学的发展,两个与代谢有关的新名词,即代谢组和代谢组学便应运而生了。代谢组(metabolome)也叫做小分子清单(small molecule inventory,SMI),是指反映细胞状态的各种小分子的样式,包括所有代谢过程(合成代谢和分解代谢)的总和以及相关的细胞过程。例如,天然和异生(xenobiotic)物质的吸收、运输和解毒,能量利用,信号转导以及调节等。如果基因组代表可能是什么,蛋白质组就代表的是表达什么,而代谢组表示的就是细胞或组织的当前状况是什么。代谢组学就是研究单个细胞或组织内所有小分子成分及其变化规律的一门学科。

实际上,代谢组给出的是某个时刻的细胞活动及其环境的直接的"画面",反映出细胞健康、疾病、衰老的状态以及药物和环境对细胞活动的影响。

代谢组学研究的一种方法是分离和定量分析构成代谢组的各种小分子化合物。例如,在使用常规的手段分离以后,可让各种小分子通过一系列电化学传感器(flow-through electrochemical sensor)进行选择性氧化或还原,在此基础上再进行测量和分析。

推荐网址:

e18-1 本章小结填填看

1. http://en.wikipedia.org/wiki/Metabolism(维基百科有关代谢的内容)
2. http://en.wikipedia.org/wiki/Metabolomics(维基百科有关代谢组学的内容)
3. http://web.expasy.org/cgi-bin/pathways/show_thumbnails.pl(一个详尽的在线代谢通路示意图)
4. http://dwb4.unl.edu/chem/chem869p/chem869plinks/www.gwu.edu/~mpb/(美国内布拉斯加大学林肯分校给出的各种代谢途径图解,包括各种代谢物的 3D 结构模型)

参考文献:

1. Nicholson J K, Lindon J C. Systems biology: Metabonomics. *Nature*, 455(7216): 1 054–1 056, 2008.
2. Milani M, et al. Substrate channeling. *Biochemistry and Molecular Biology Education*, 31(4): 228–233, 2003.
3. James C L, Viola R E. Production and characterization of bifunctional enzymes. substrate channeling in the aspartate pathway. *Biochemistry*, 41: 3 726–3 731, 2002.
4. Giersch C. Mathematical modeling of metabolism. *Curr Opin Plant Biol*, 3: 249–253, 2000.

第十九章　生物能学

生物能学(bioenergetics)也称为生化热力学(biochemical thermodynamics),它专门研究生命系统内能量流动和转化的规律。通过这门分支学科,可以更好地理解和解释生命现象之中很多与能量有关的问题。

本章将主要介绍热力学定律、自由能变化与生化反应的关系、生物体内的偶联反应以及高能生物分子的结构与功能。

第一节　热力学定律与 Gibbs-Helmholtz 方程

为了更好地理解生物能学,首先要认识热力学的两个定律:热力学第一定律即能量守恒定律;热力学第二定律是指一个系统的熵倾向于增加,或者一个系统做功的能力趋于下降。其次要有所了解与热力学相关的几个基本概念。

(1) 系统(system)和环境(surrounding)　系统为宇宙之中人们感兴趣的任何对象,环境是指一个系统周围的任何事物。系统又可以分为孤立系统(insulated system)、封闭系统(closed system)和开放系统(open system)。其中,孤立系统与环境之间不发生任何形式的物质交换和能量交换,封闭系统则可以与环境之间进行能量的交换,开放系统和环境之间既可以进行能量交换也可以进行物质交换(图 19-1)。显然,生命系统属于开放系统。

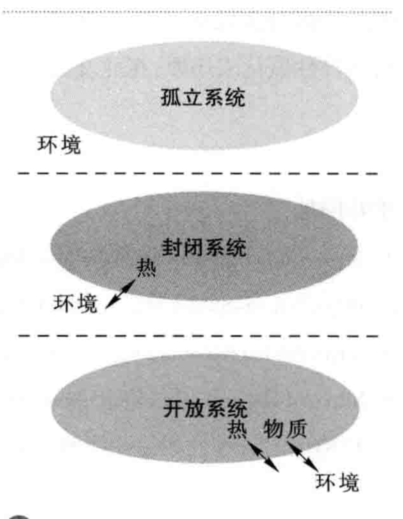


图 19-1　孤立系统、封闭系统和开放系统

(2) 能量　是指做功的本领。

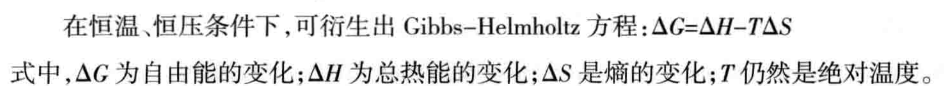
(3) 自由能(free energy)　是指一个系统的总能量之中用来做功的那一部分能量,即有用能。

(4) 熵(entropy)　是指一个系统的无序状态。一个系统越有序,它的熵就越低。

(5) 焓(enthalpy)　是指系统的总热能。

综合热力学第一定律和第二定律即可得出以下方程:$G=H-T-S$

其中,G 代表系统的自由能;S 代表系统的熵;H 为系统的总热能;T 为系统的绝对温度。上述方程适合任何系统。



在恒温、恒压条件下,可衍生出 Gibbs-Helmholtz 方程:$\Delta G=\Delta H-T\Delta S$

式中,ΔG 为自由能的变化;ΔH 为总热能的变化;ΔS 是熵的变化;T 仍然是绝对温度。

由于多数生命系统处于恒温(37℃)和恒压(101.3kPa)条件下,因此上述方程同样适用于生命系统,包括发生在体内的每一个生化反应。



第二节　生化反应的方向性与自由能之间的关系

假设一个生化反应:$A+B \longleftrightarrow C+D$

在恒温和恒压下,不难推导出以下公式(推导过程从略):$\Delta G=\Delta G^{\ominus\prime}+RT\ln\dfrac{[C][D]}{[A][B]}$

式中,ΔG 为反应的自由能变化;$\Delta G^{\ominus\prime}$ 为在标准条件下(反应温度为 25℃,压强为 101.3 kPa,底物和产物的浓度均为 1 mol/L,pH 为 7)的自由能变化;R 为气体常数(为 8.31×10^{-3} kJ·mol^{-1}·K^{-1},K 为玻尔兹曼常数)。

Quiz1 生化反应的标准条件与纯化学反应的标准条件有何不同?

以上公式显示,一个生化反应的自由能变化包括两个部分:一部分是恒定的,它由底物和产物固有的性质决定;另一部分是可变的,它由反应的温度、底物与产物的浓度决定。在给定的条件下,通过上述公式可以计算出一个反应的自由能的变化值,但必须首先确定在标准条件下的自由能变化即 $\Delta G^{\ominus\prime}$。

如果反应到达平衡,则 ΔG=0。上面的公式可作以下的转变(K'_{eq} 为反应的平衡常数):

$$\Delta G=\Delta G^{\ominus\prime}+RT\ln\frac{[C][D]}{[A][B]}\longrightarrow 0=\Delta G^{\ominus\prime}+RT\ln\frac{[C][D]}{[A][B]}\longrightarrow \Delta G^{\ominus\prime}=-RT\ln\frac{[C][D]}{[A][B]}$$

$$\text{已知 } K'_{eq}=\frac{[C][D]}{[A][B]},\text{则 } \Delta G^{\ominus\prime}=-RT\ln K'_{eq} \text{ 或 } \Delta G^{\ominus\prime}=-2.303RT\lg K'_{eq}$$

因此,如果一个反应的平衡常数已知,则很容易通过上式计算出该反应的标准自由能变化。

一个反应的 $\Delta G^{\ominus\prime}$确定以后,那么此反应在一个给定的非标准条件下的 ΔG 很容易计算出来。通过计算 ΔG,可以判断一个反应在给定条件下进行的方向。

如果 $\Delta G<0$,该反应就可以自发地进行,此反应为放能反应(exergonic)或下坡反应(downhill reaction)。需要特别注意的是,如果 ΔG 为一个非常大的负值,则表明此反应趋于完全,为不可逆反应;如果 ΔG=0,反应就处于平衡状态,底物和产物的浓度维持不变;如果 $\Delta G>0$,此反应就不能自发地进行,除非向此反应提供能量,因此该反应为需能反应(endergonic)或上坡反应(uphill reaction)。如果 ΔG 是一个非常大的正值,则意味着整个底物处于一种非常稳定的状态,反应几乎没有发生的可能。

以糖酵解中的磷酸二羟丙酮异构成 3- 磷酸甘油醛的反应为例,这步反应的平衡常数为 0.475,将它代入 $\Delta G^{\ominus\prime}=-2.303RT\lg K'_{eq}$,则得:

$$\Delta G^{\ominus\prime}=-2.303\times8.314\times10^{-3}\times310\times\lg0.475=1.92\ \text{kJ/mol}。$$

$\Delta G^{\ominus\prime}$ 为正值,说明此反应在标准条件下是不利的。但细胞内的[磷酸二羟丙酮]=2×10^{-4} mol/L,[3- 磷酸甘油醛]=3×10^{-6} mol/L,因而:

$$\Delta G=\Delta G^{\ominus\prime}+RT\ln\frac{[\text{3- 磷酸甘油醛}]}{[\text{磷酸二羟丙酮}]}=1.92+8.314\times10^{-3}\times310\times\ln\frac{3\times10^{-6}}{2\times10^{-4}}=-8.9\ \text{kJ/mol}$$

这表明在细胞内,此反应仍可以自发地进行。

由此可见,一个 $\Delta G^{\ominus\prime}$为正值的反应在细胞内仍然可以朝正反应方向进行,只要底物浓度足够高,或者产物的浓度足够低。

需要特别注意的是,任何在热力学上有利,可以自发进行的反应,都具有活化能的障碍需要克服。热力学定律只能预测一个反应能否自发进行,但并不能预测出一个反应的动力性性质,即一个反应有没有发生以及反应速率是多少这种现实性。例如,葡萄糖氧化成 CO_2 和 H_2O 的反应,其 $\Delta G^{\ominus\prime}$为 −5 693 kJ/mol,是一个非常大的负值。这意味着葡萄糖的氧化在热力学上是一个极为有利的反应,但这个反应如果没有酶的催化或者直接将其用火燃烧的话,糖罐子里的葡萄糖在室温下很难被氧化成 H_2O 和 CO_2。

Quiz2 根据热力学第二定律解释:金刚石转变为石墨是有利的,但却很难发生。

第三节　ΔG 与 ΔE 之间的关系

细胞内的很多反应为氧化还原反应,一个氧化还原反应的 $\Delta G^{\ominus\prime}$ 与其标准氧化还原电位(标准条件与自由能的标准条件一样)的变化值 $\Delta E^{\ominus\prime}$可用公式 $\Delta G^{\ominus\prime}=-nF\Delta E^{\ominus\prime}$表示,$n$ 为电子转移数,F 是法拉第(Farady)常数,为 96.485 kJ·mol^{-1}·V^{-1},$\Delta E^{\ominus\prime}$ = 含有氧化剂或电子受体的半反应 $E^{\ominus\prime}$ − 含有还原剂或电子供体的半反应的 $E^{\ominus\prime}$。而非标准条件下的 ΔG 与同样条件下的 ΔE 之间的关系为:$\Delta G=-nF\Delta E$。

假设一个氧化还原反应 $A+BH_2 \longrightarrow AH_2+B$，此反应实际上由两个半反应即还原半反应(reductive half-reaction) $A+2H^++2e^- \longrightarrow AH_2$ 和氧化半反应(oxidative half-reaction) $BH_2 \longrightarrow B+2H^++2e^-$ 组成，根据 Nernst 方程：

$$\Delta E=\Delta E^{\ominus\prime}-\frac{2.3RT}{nF}\lg\frac{[AH_2][B]}{[A][BH_2]}，则\ \Delta G=-nF\left(\Delta E^{\ominus\prime}-\frac{2.3RT}{nF}\lg\frac{[AH_2][B]}{[A][BH_2]}\right)$$

$$=2.3RT\lg\frac{[AH_2][B]}{[A][BH_2]}-nF\Delta E^{\ominus\prime}=2.3RT\lg\frac{[AH_2][B]}{[A][BH_2]}+\Delta G^{\ominus\prime}$$

下面以丙酮酸 $+NADH+H^+ \longrightarrow$ 乳酸 $+NAD^+$ 反应为例加以说明。该反应由两个半反应：半反应 1，丙酮酸 $+2H^++2e^- \longrightarrow$ 乳酸；半反应 2，$NADH+H^+ \longrightarrow NAD^++2H^++2e^-$ 组成，其中半反应 1 的 $E^{\ominus\prime}=-0.190$ V，半反应 2 的 $E^{\ominus\prime}=-0.320$ V，由于丙酮酸为氧化剂(电子受体)，NADH 为还原剂(电子供体)，所以，$\Delta E^{\ominus\prime}=-0.190\ V-(-0.320\ V)=0.130\ V$，而 $\Delta G^{\ominus\prime}=-nF\Delta E^{\ominus\prime}=-2\times96.48\times0.130=-25.1$ kJ/mol。

$\Delta G^{\ominus\prime}$ 为负值说明在标准条件下有利于正反应(乳酸的形成)的进行。

如果不是在标准条件下，而是反应混合物中[乳酸]/[丙酮酸]和[NAD^+]/[NADH]均为 1 000 : 1，则 $\Delta G=2.3RT\lg\frac{[AH_2][B]}{[A][BH_2]}-nF\Delta E^{\ominus\prime}=2.303\times8.314\times10^{-3}\times310\times\lg(1\ 000\times1\ 000)-25.1=10.5$ kJ/mol。此时 ΔG 为正值，这说明在此条件下有利于逆反应的进行。

第四节　生命系统内的偶联反应

假定一个反应 A → B，$\Delta G_1=G_B-G_A>0$，则此反应不能自发地进行，但如果此反应通过某种机制与另外一个 $\Delta G_2=G_D-G_C<0$ 的反应 C → D 偶联在一起，而且总的自由能变化 $\Delta G=\Delta G_1+\Delta G_2<0$，则 A → B 的反应照样可以自发进行，在这里第二个反应释放出的能量被用来驱动第一个反应。

在生物体内，有两种偶联机制。第一种机制通过一个共同的代谢中间物来实现(图 19-2)，反应式可简写为：A+C → I → B+D。

以葡萄糖的磷酸化反应为例加以说明：葡萄糖 +Pi $\longrightarrow$ 6- 磷酸葡糖 $+H_2O$。

该反应的 $\Delta G^{\ominus\prime}=13.8$ kJ/mol，显然这样的反应在热动力学上是极端不利的。再来看另外一个反应：$ATP+H_2O \rightarrow ADP+Pi$。此反应的 $\Delta G^{\ominus\prime}=-30.5$ kJ/mol，是一个极大的负值。如果这两个反应能够偶联起来，总反应式就为：葡萄糖 +ATP $\longrightarrow$ 6- 磷酸葡糖 +ADP。

总的自由能变化 $\Delta G^{\ominus\prime}_{总}=13.8+(-30.5)=-16.7$ kJ/mol，依然是一个较大的负值，也就意味着葡萄糖可顺利地磷酸化为 6- 磷酸葡糖。实际上在细胞内，在己糖激酶或葡萄糖激酶的催化下，这两个反应正是紧密地偶联在一起的。

偶联反应的第二种机制是通过特殊的高能生物分子(high-energy biomolecule)来进行的：其中在第一个反应释放出的自由能中，有一部分转变为高能生物分子(如 ATP)之中的化学能贮存起来，而第二个反应的进行则由第一个反应形成的高能生物分子来驱动。细胞内有很多这样的例子，一般情况是细胞内的分解代谢产生生物高能分子，合成反应则利用这些高能分子(图 19-2)。

偶联机制一：

A
放能反应
D
热
自由能
化学能
需能反应
C
B
A+C ⟶ B+D+ 热

偶联机制二：

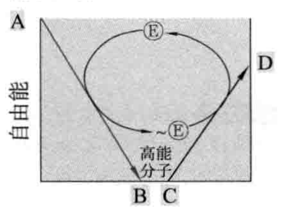


图 19-2　放能反应与吸能反应偶联的两种机制

第五节　高能生物分子

高能生物分子又简称为高能分子，它是指那些既容易水解又能够在水解之中释放出大量自由能($\Delta G^{\ominus\prime}$ 为极大的负值)的一类分子的总称，以高能磷酸化合物(high-energy phosphate compound)最为常

见。在高能分子水解的时候，被水解断裂的化学键似乎贮存着大量的能量，因此有人称此键为高能键(high-energy bond)，经常用“~”表示。

► 表 19-1 常见的高能分子

名称	高能键水解以后的产物	高能键水解释放出的 $\Delta G^{\ominus\prime}$/(kJ·mol^{-1})
磷酸烯醇式丙酮酸	丙酮酸 +Pi	–62
3',5'-cAMP	5'-cAMP	–50.4
1,3- 二磷酸甘油酸	3- 磷酸甘油酸 +Pi	–49.6
磷酸肌酸	肌酸 +Pi	–43.3
乙酰磷酸	乙酸 +Pi	–43.3
ATP	ADP+Pi	–35.7
ATP	AMP+PPi	–35.7
PPi	2Pi	–33.6
尿苷二磷酸葡糖	UDP+ 葡萄糖	–31.9
乙酰 -CoA	乙酸 +CoA	–31.5
S- 腺苷甲硫氨酸	腺苷 +Met	–25.6

表 19-1 为一些常见的高能分子的名称、结构和相应的 $\Delta G^{\ominus\prime}$ 值，其中磷酸烯醇式丙酮酸“高高在上”，其水解时释放的能量最高，大约为 ATP 的两倍。而 ATP 正好位于“中游”的位置，这意味着它合成起来容易，利用起来也容易。正是 ATP 在众多的高能分子之中所处的独特位置，使它成为最重要的能量载体。通过它，细胞内的放能反应和需能反应很容易发生偶联。

Quiz3 在以下生物分子——cGMP、辅酶 I、辅酶 II、FAD、琥珀酰 -CoA、2,3-BPG、ADP、6- 磷酸果糖中，哪些含有高能键?

以糖酵解最后一步合成 ATP 的反应(参看第二十二章“糖酵解”)为例，此反应的 $\Delta G^{\ominus\prime}$ =–31.3 kJ/mol，磷酸烯醇式丙酮酸为底物。糖酵解的第一步反应则是一步消耗 ATP 的反应，$\Delta G^{\ominus\prime}$ =–16.7 kJ/mol。这说明一方面葡萄糖分解释放出的一部分能量驱动了 ATP 合成，另一方面新合成的 ATP 又可以反过来激活新的葡萄糖分子，而使得更多的葡萄糖分子进入糖酵解。那么，磷酸烯醇式丙酮酸能否直接将其磷酸基团转移到葡萄糖，而不需要 ATP 作为中介物呢? 答案是不能，因为细胞内根本就不存在相关的酶。因此，没有 ATP 的合成，细胞内各种需能反应是不可能正常进行的。

在细胞内，作为能量载体的 ATP 几乎参与所有的生理过程(图 19-3)，例如肌肉收缩、生物合成、细胞运动、细胞分裂、主动转运、神经传导等，因此有人称之为通用的“能量货币”(energy currency)。细胞的正常生理活动需要消耗大量的 ATP，所以细胞内 ATP 的周转率(turnover rate)非常高。据估计，一个人在 24 h 内水解及再合成的 ATP 总量相当于他的体重。ATP 如此高的周转率使得它并不适合充当能量的贮存者。例如，肌肉运动时需要消耗大量的 ATP，以至于水解 ATP 的速率远远大于重新合成 ATP 的速率，在生物进化的过程中，磷酸肌酸的出现解决了 ATP 的供需矛盾。在静息状态下，ATP 浓度较高，磷酸肌酸作为“黄金储备”得以大量合成，以备不时之需。

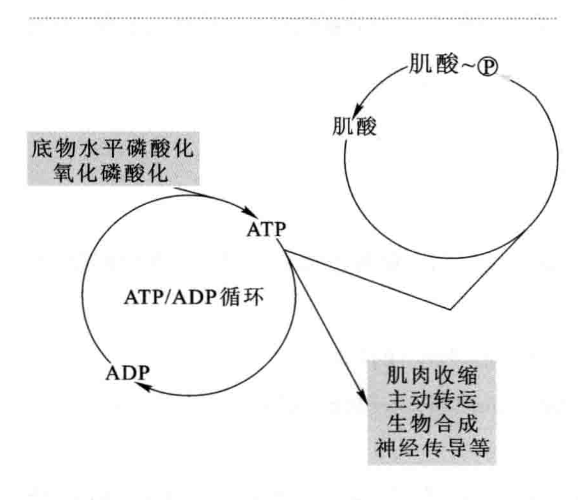


图 19-3 ATP 的合成和利用

磷酸肌酸并不是唯一的贮能物质，某些生物(如龙虾和蟹类)还可以使用磷酸精氨酸代替磷酸肌酸，而某些微生物则使用聚偏磷酸(polymetaphosphate)作为贮能物质(图 19-4)。一般将磷酸肌酸和磷酸精氨酸通称为磷酸原(phosphagen)。当 ATP 水解为 ADP 和无机磷酸之后，尽管 ADP 也是高能分子，但机体很少直接使用它作为能量供体，而是通过腺苷酸激酶(adenylate kinase)将 ADP 转变为 ATP。

ATP 的利用方式除了将它转变为 ADP 以外，还可以将它转变为 AMP 和焦磷酸，由于焦磷酸很容

磷酸肌酸　　磷酸精氨酸　　聚偏磷酸

图 19-4　几种贮能分子的化学结构

易被细胞内的焦磷酸酶水解成为两个无机磷酸，致使总的 $\Delta G^{\ominus\prime}$ 更负（-27.6 kJ/mol）。例如脂肪酸的活化反应：ATP+CoASH+RCOOH ⟶ AMP+PPi+RCO-SCoA（详见第二十九章“脂肪酸代谢”）。焦磷酸的水解不仅可以使得反应更加完全，也使得磷酸根可以循环利用。

除了 ATP 以外，其他 3 种核苷三磷酸有时也可以作为能量货币，这几种能量货币在细胞内是可以自由“兑换”的，但需要核苷二磷酸激酶（nucleotide diphosphate kinase）的催化（参看第三十一章“氨基酸代谢”）。

$$\text{ATP+NDP} \xrightleftharpoons{\text{核苷二磷酸激酶}} \text{ADP+NTP}$$

$$\text{ATP+AMP} \xrightleftharpoons{\text{腺苷酸激酶}} \text{2ADP}$$

ATP 除了作为“能量货币”使用以外，还可以作为很多酶的别构效应物来调节相关酶的活性。

高能分子在水解时能释放出大量的能量（以 ATP 为例）主要与两个因素有关：①相对 ATP 而言，水解的产物 ADP 和无机磷酸具有更大的共振稳定性（resonance stability）（图 19-5）；②水解以后，负电荷之间的静电斥力减弱了。

等等

图 19-5　ATP 及其水解物的共振结构

细胞内 ATP 的合成就是 ADP 被磷酸化形成 ATP 的过程，机体有 3 种合成 ATP 的手段：底物水平磷酸化（substrate-level phosphorylation）（参看第二十二章“糖酵解”）、氧化磷酸化（详见第二十章“生物氧化”）和光合磷酸化（详见第二十六章“光合作用”）。

e19-1 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/thermodynamics.html（完全免费的医学生物化学课程网站，有关热力学的内容）
2. http://en.wikipedia.org/wiki/Bioenergetics（维基百科有关生物能学的内容）
3. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb1/part2/bioener.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关生物能学的内容）
4. http://www.bmb.leeds.ac.uk/illingworth/oxphos/physchem.htm（英国利兹大学生物系提供的与生物能学有关的内容）

参考文献：

1. Green D E, Zande H D. Universal energy principle of biological systems and the unity of bioenergetics. *Proc Natl Acad Sci USA*, 78(9): 5 344-5 347, 1981.
2. Stuart J, et al. Bioenergetics. 3rd ed. New York: Academic Press, 2002.

第二十章　生物氧化

生物体内发生的氧化反应统称为生物氧化。与体外发生的非生物氧化共同点是：反应的本质都是脱氢、失电子或加氧；被氧化的物质相同，终产物和释放的能量也相同。在需氧生物体内，糖类、脂肪和蛋白质等代谢物经生物氧化的终产物均为 H_2O 和 CO_2，其中 CO_2 由生物氧化中形成的含有羧基的中间物经脱羧反应产生。H_2O 的形成比较复杂，一般是由代谢物脱下来的氢经电子传递体先形成质子，失去的电子经一系列电子传递体传给 O_2 形成氧负离子，最后质子与氧负离子结合生成 H_2O。

与非生物氧化不同的是：生物氧化的主要方式为脱氢；生物氧化是在酶的催化下进行，因此条件比较温和；生物氧化是在一系列酶、辅酶（辅基）和电子传递体的作用下逐步进行的，每一步释放一部分能量。以葡萄糖为例，它在细胞内需要经历几十步反应才能最终氧化成 CO_2 和 H_2O。而在细胞外的非生物氧化即燃烧几乎是“一蹴而就”的，一把火可在顷刻间将其氧化成 CO_2 和 H_2O。生物氧化所具有的这一特征的优点在于：既可以防止能量的骤然释放而损害有机体，又可以让机体更好地捕获能量合成 ATP。

本章将主要介绍在生物氧化过程中呼吸链上的电子传递和氧化磷酸化的机制，特别是“化学渗透”学说和“结合变化”学说会重点阐述。

第一节　呼吸链

生物氧化过程中，从代谢物脱下来的高能电子需要经过一系列中间传递体，最后才能交给氧气，在其间能量逐步释放。这种由一系列电子传递体构成的链状复合体称为电子传递体系（electron transport system，ETS），或者简称为呼吸链（respiratory chain）。按照最初的电子来源，呼吸链一般可分为 NAD^+ 呼吸链和 FAD 呼吸链。两种呼吸链位于原核细胞的细胞膜和真核细胞的线粒体内膜上，主要功能是通过与氧化磷酸化的偶联产生 ATP。

一、呼吸链的组分

呼吸链的主要成分如下：

(1) NAD^+ 及与 NAD^+ 偶联的脱氢酶　此类电子传递体以 NAD^+ 为辅酶或底物。其传递电子或氢的机制是：先是 NAD^+ 与一个脱氢酶的活性中心结合，从代谢物得到氢和电子以后被还原成 NADH/H^+，随后 NADH 离开酶的活性中心，扩散到呼吸链的“入口”，留下氢和电子后离开呼吸链去再次与脱氢酶结合，并重复先前的过程。由此可见，在传递电子的过程中，NAD^+ 不断往返于代谢物和呼吸链之间，这种性质使之成为一种流动的电子传递体（mobile electron carrier）。

(2) 黄素及与黄素偶联的脱氢酶　此类电子传递体以 FMN 或 FAD 作为辅基，两个典型的代表是 NADH 脱氢酶和琥珀酸脱氢酶。

FMN 或 FAD 能够作为电子传递体是因为它们有氧化型和还原型两种形式（图 20–1）。氧化型的 FMN 或 FAD 可以接受两个电子和两个质子，而成为还原型的 $FMNH_2$ 或 $FADH_2$。具体的传递过程可以分两步进行，也可以一步到位。如果是一步到位，就两个电子同时传递，没有半醌中间物。如果分两步，每一步传递一个电子，中间有一个半醌中间物（FMNH·或 FADH·）。这是一种反应性极强的自由基，与水相遇可能会丢掉电子，启动自由基链式反应，因此在作为脱氢酶的辅助因子的时候，黄素总是以辅基

的形式与酶分子紧密结合，有时甚至与之共价结合。还因为 FAD 在酶活性中心，与多个氨基酸侧链上的功能基团有作用，因此与酶结合的 FAD 的 $E^{\ominus\prime}$ 值变化很大，有时比 NAD^+ 大，有时则小于 NAD^+。

图 20-1 FMN 和 FAD 参与的电子传递

Quiz1 机体有哪些天然的化合物，可作为抗氧化剂帮助机体清除自由基?

（3）辅酶 Q（coenzyme Q，CoQ） CoQ 是一种脂溶性的醌类化合物，其结构与质体醌相似，因其广泛存在于自然界，故又名为泛醌（ubiquinone，UQ）。不同来源的 CoQ 的侧链长度不一定相同，这与异戊二烯单位的数目（n 值）有关。n 值一般为 6~10，其中哺乳动物体的 n 值为 10，故它的 CoQ 常被简称为 Q_{10}。CoQ 主要存在于线粒体内膜上，作为呼吸链的重要成分参与电子的传递，其也存在其他的生物膜上，作为一种抗氧化剂，有利于机体清除自由基。

CoQ 之所以能够充当电子传递体，是因为它具有氧化型和还原型两种形式。在细胞内，这两种形式是可以相互转变的（图 20-2）：氧化型的 CoQ 可以接受两个质子和两个电子，形成还原型的 $CoQH_2$。而还原型的 $CoQH_2$ 可以失去两个质子和两个电子，重新转变为氧化型的 CoQ。与 FMN 和 FAD 一样，辅酶 Q 的得（失）电子既可以分两步进行［一次得（失）一个电子］，也可以一步到位［同时得（失）两个电子］。这一点非常重要，因为在呼吸链上某些电子传递体（单电子传递体）一次只能得（失）

图 20-2 辅酶 Q 的化学结构及其电子的传递功能

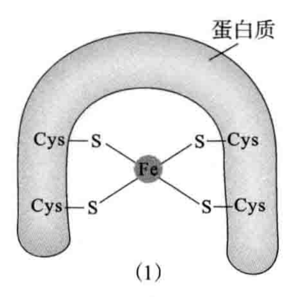
一个电子，而某些电子传递体（双电子传递体）一次必须得（失）两个电子。CoQ 在两者之间正好起着一种过渡的作用，即它从双电子传递体一次接受两个电子，再分两次传给单电子传递体。

CoQ 在线粒体内膜上的含量远远超过呼吸链上的其他成员，其脂溶性的性质使得它在膜上具有高度的流动性，因此它特别适合作为一种流动的电子传递体，在两个非流动的电子传递体之间传递电子。

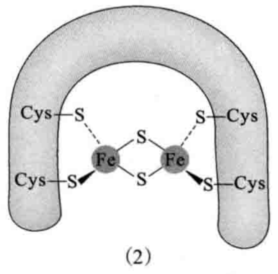
(4) 铁硫蛋白（iron-sulfur protein）又称为铁硫中心（iron-sulfur center）或铁硫簇（iron-sulfur cluster） 其分子之中含有非血红素铁，还可能含有对酸不稳定的无机 S，其中 Fe 通过配位键与含有孤对电子的硫结合。铁硫蛋白是借助于铁的价态变化来传递电子。

Quiz2 如何确定一种铁硫蛋白有无无机硫？

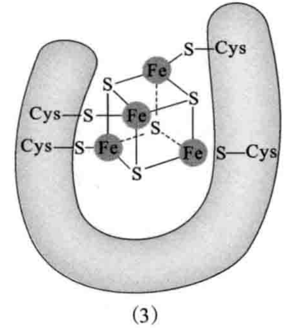
已知铁硫蛋白有 3 类（图 20–3）：第一类含有单个 Fe，但没有无机 S。此类铁硫蛋白上的单个 Fe 与 4 个 Cys 残基上的巯基 S 相连；第二类含有 2 个 Fe 和 2 个无机 S（Fe_2S_2），其中每个 Fe 各与 2 个无机 S 和 2 个 Cys 残基的巯基 S 相连；第三类含有 4 个 Fe 和 4 个无机 S（Fe_4S_4），其中的 Fe 与 S 相间排列在一个正六面体的 8 个顶点，此外，4 个 Fe 还各与 1 个 Cys 残基上的巯基 S 相连。



(1)



(2)



(3)

图 20–3 三类铁硫蛋白的结构

(5) 细胞色素 它也是借助于铁的价态变化来传递电子，但与铁硫蛋白不同，细胞色素含有血红素铁。

细胞色素广泛存在于需氧生物。目前已发现的细胞色素有 30 多种，每一种细胞色素都有特殊的光吸收。它们的还原形式通常具有 α、β 和 γ 三个吸收峰。其中 α 吸收峰的差别最大，因此有人根据此吸收峰的波长，将细胞色素分为 a、b 和 c 三大类，与它们结合的血红素辅基分别称为血红素 a、b 和 c。每一大类还可以细分。例如，a 类细胞色素可分为细胞色素 a 和 a_3，这两种细胞色素通常形成紧密的复合物。b 类细胞色素可分为细胞色素 b_{562} 和细胞色素 b_{566} 等，c 类细胞色素可分为细胞色素 c 和 c_1。

细胞色素 aa_3 上的铁离子只形成了 5 个配位键，因此还能与 O_2 再形成一个配位键，从而将上游传过来的电子直接交给氧气。但铁离子也可与 CO、氰化物、H_2S 或叠氮化物等含有孤对电子的物质形成配位键，因而如果细胞色素 aa_3 与 O_2 以外的物质（如氰化物）结合，就会阻断整个呼吸链的电子传递，引起机体中毒。

(6) 分子氧。

二、呼吸链组分的排列顺序

不论是在原核细胞的质膜上，还是在真核细胞的线粒体内膜上，构成呼吸链的各个组分都是按照一定的次序排列的。使用下述几种方法，人们可以确定它们在呼吸链上的排列顺序。

(1) 测定各成分的标准氧化还原电位（$E^{\ominus\prime}$） $E^{\ominus\prime}$值表示一种物质的氧化还原能力，$E^{\ominus\prime}$值越小，其还原性越强，更容易被氧化；$E^{\ominus\prime}$值越大，其氧化性越强，更容易被还原（表 20–1）。因此，呼吸链中各种组分的排列顺序应当由低 $E^{\ominus\prime}$依次向高 $E^{\ominus\prime}$排列。但是，由于各传递体在体外测得的 $E^{\ominus\prime}$值与它们在内膜环境上实际的 $E^{\ominus\prime}$值可能有出入，因此，对于此方法得出的结论必须谨慎对待。

Quiz3 治疗氰化物中毒的一种方法是让中毒者服用适量的亚硝酸盐和硫代硫酸钠，你认为其中的生化机理是什么？

► 表 20–1 电子传递体的标准氧化还原电位

还原反应（半反应）	$E^{\ominus\prime}$/V
$2H^+ + 2e^- \rightarrow H_2$	–0.414
$NAD^+ + H^+ + 2e^- \rightarrow NADH+$	–0.320
NADH 脱氢酶（FMN）+ $2H^+ + 2e^- \rightarrow$ NADH 脱氢酶（$FMNH_2$）	–0.30
$CoQ + 2H^+ + 2e^- \rightarrow CoQH_2$	0.045
细胞色素 b（Fe^{3+}）+ $e^- \rightarrow$ 细胞色素 b（Fe^{2+}）	0.077
细胞色素 c（Fe^{3+}）+ $e^- \rightarrow$ 细胞色素 c（Fe^{2+}）	0.22
细胞色素 c_1（Fe^{3+}）+ $e^- \rightarrow$ 细胞色素 c_1（Fe^{2+}）	0.254
细胞色素 a（Fe^{3+}）+ $e^- \rightarrow$ 细胞色素 a（Fe^{2+}）	0.29
细胞色素 a_3（Fe^{3+}）+ $e^- \rightarrow$ 细胞色素 a_3（Fe^{2+}）	0.55
$1/2O_2 + 2H^+ + 2e^- \rightarrow H_2O$	0.816

(2) 确定在有氧环境下氧化反应达到平衡时各电子传递体的还原程度　在有氧条件下，测定离体线粒体的三羧酸循环反应达到平衡时，呼吸链中各组分的还原程度。倘若反应达到平衡，则从呼吸链的起点到终端，各组分的还原程度一定是递减的，即越靠近氧气，其还原程度就越低。这种情况类似于物理学上的联通管，水流相当于电子流，水位相当于还原程度。若进水量 = 出水量(流量达到平衡)，则水位从进水管到出水管逐渐减低，离进水口最近的水管中的水位最高，而离出水管最近的水管中的水位最低。

鱼藤酮

NADH ⊗→ Q → Cyt b → Cyt c_1 → Cyt c → Cyt ($a+a_3$) → O_2

抗霉素A

NADH → Q → Cyt b ⊗→ Cyt c_1 → Cyt c → Cyt ($a+a_3$) → O_2

CN^-、CO、H_2S或叠氮化物

NADH → Q → Cyt b → Cyt c_1 → Cyt c → Cyt ($a+a_3$) ⊗→ O_2

图 20-4　几种呼吸链抑制剂的化学结构及其作用位点

(3) 使用特异性呼吸链抑制剂和人工电子受体　使用位点特异性抑制剂阻断呼吸链中的特定环节，经过一定时间以后，在被阻断部位的上游，各个传递体应为还原型，而下游一侧的应为氧化型，这正像用塞子在某一处将联通管堵住一样，受堵部位前面各水管中的水应该是满的(全是还原型)，而受堵部位后面各水管中的水很快就流光(全是氧化型)。然而，假定在堵塞部位的前方开一个小孔，则水流将“另辟蹊径”，恢复流动。人工电子受体(artificial electron acceptor)(如铁氰化钾和亚甲蓝)的作用就相当于上述的小孔，它可以插入到呼吸链上特定的部位，接受 $E^{\ominus\prime}$ 值比它低的传递体上的电子，使得电子恢复流动。因此，通过分析不同的抑制剂和不同的人工电子受体作用呼吸链的结果，也可以确定呼吸链中各传递体的排列顺序。

图 20-4 为几种常见的呼吸链抑制剂的作用位点，显然，它们对于绝大多数生物都是有毒的。例如，氰化钠有剧毒，鱼藤酮曾经被渔民用做捕杀鱼类的毒饵，将少量的叠氮化物放入存放树脂溶液之中，可防止微生物的污染。

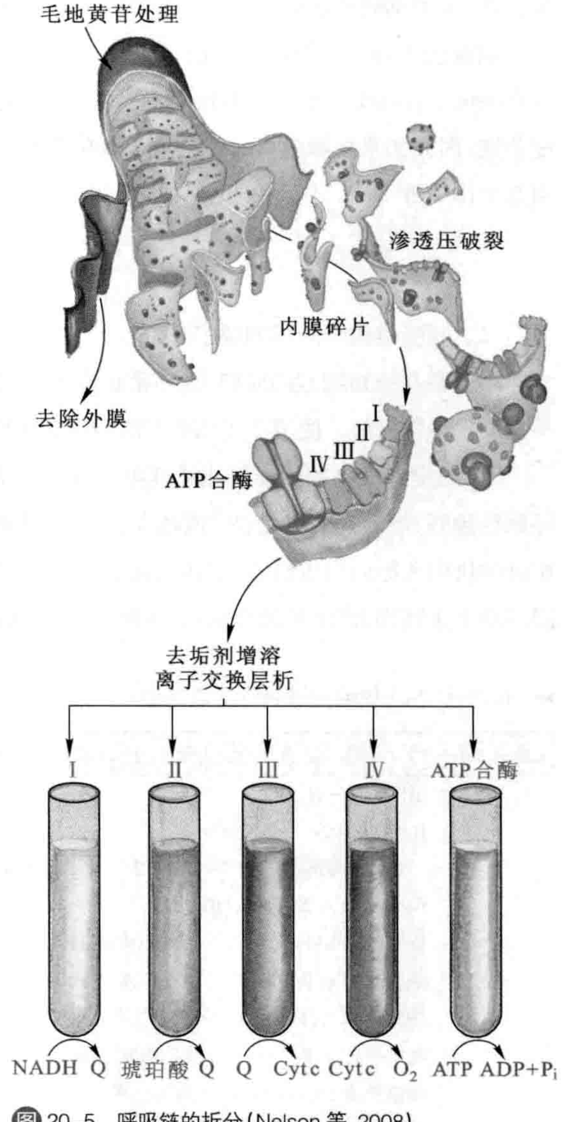
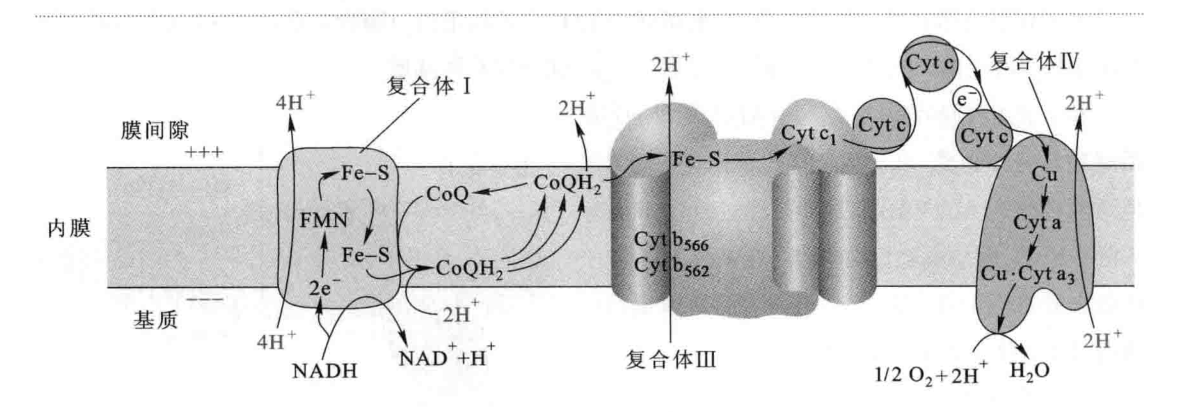


图 20-5　呼吸链的拆分(Nelson 等，2008)

(4) 呼吸链的拆分和重组　使用图 20-5 所示的方法，可将呼吸链拆分成 4 种具有不同催化活性的复合物(复合体Ⅰ~复合体Ⅳ)以及游离的 CoQ 和细胞色素 c。通过研究各复合物

的结构和组成，以及它们在体外所能催化反应的性质，同时结合其他几种方法，已基本确定呼吸链上各组分的精确排列顺序(图 20-6)。从图中可以看出，呼吸链是以复合体的形式组织在一起的。其中 NAD^+ 呼吸链由复合体Ⅰ、Ⅲ、Ⅳ以及两种流动的电子传递体 CoQ 和细胞色素 c 共同组成，FAD 呼吸链则由复合体Ⅱ、Ⅲ、Ⅳ以及同样的两种流动的电子传递体共同组成。CoQ 在复合体Ⅰ和Ⅲ或复合体Ⅱ和Ⅲ之间传递电子，细胞色素 c 在复合体Ⅲ和Ⅳ之间传递电子。电子到达每一个复合体以后，先沿着内部的环路进行传递，再通过流动的电子传递体传给下一个复合体。NAD^+ 呼吸链电子传递的方向是：复合体Ⅰ→ CoQ →复合体Ⅲ→细胞色素 c →复合体Ⅳ→ O_2。FAD 呼吸链电子传递的方向是：复合体Ⅱ→ CoQ →复合体Ⅲ→细胞色素 c →复合体Ⅳ→ O_2。



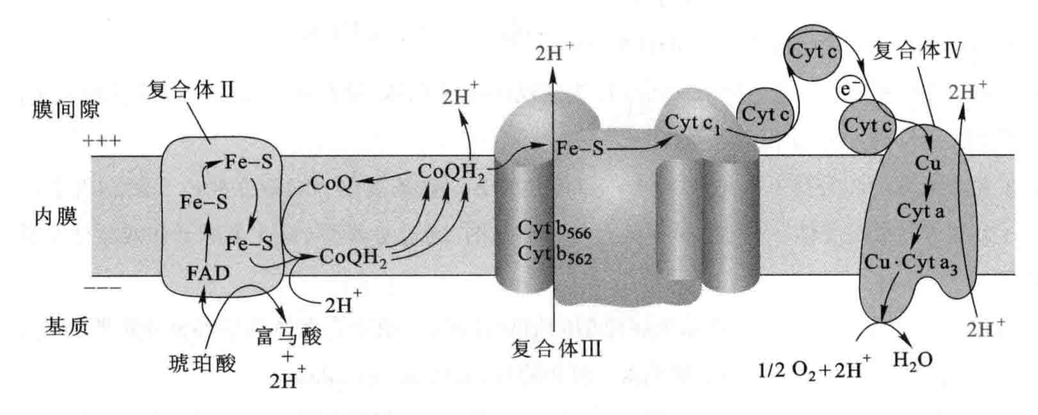


图 20-6 电子在各复合体和复合体之间的传递

三、复合体Ⅰ、Ⅱ、Ⅲ和Ⅳ的结构与功能

表 20-2 概括并比较了 4 种复合体最重要的性质，以及它们在呼吸链上的相互关系。为了进一步说明它们的结构与功能，下面就分别作介绍。

Quiz4 同等剂量的鱼藤酮、萎锈灵和抗霉素 A，你认为哪一种抑制呼吸链的效果最强？为什么？

► 表 20-2 复合体Ⅰ、Ⅱ、Ⅲ和Ⅳ的结构和性质

复合体	别名	M_r ($\times 10^6$)	多肽链的数目	辅酶或辅基	电子流动方向	一对电子产生的质子	抑制剂	在内膜上的相对比率
Ⅰ	NADH-CoQ 氧化还原酶	0.7~0.9	43，其中有 7 条由 mtDNA 编码	1 个 FMN、6~9 个铁硫蛋白	NADH → CoQ	4 个	鱼藤酮、安米妥、杀粉菌素	1
Ⅱ	琥珀酸 -CoQ 氧化还原酶	0.14	4~5	1 个 FAD、3 个铁硫蛋白	琥珀酸 → CoQ	0	萎锈灵 (carboxin)	2
Ⅲ	CoQ- 细胞色素 c 氧化还原酶	0.25	11，其中有 1 条由 mtDNA 编码	2 个血红素 b、1 个血红素 c_1、1 个铁硫蛋白	CoQ →细胞色素 c	4 个	抗霉素 A	3
Ⅳ	细胞色素 c 氧化酶	0.16~0.17	13，其中有 4 条由 mtDNA 编码	2 个 Cu、血红素 a、血红素 a_3	细胞色素 c → O_2	2 个	CO、H_2S、氰化物、叠氮化物	6~7

1. 复合体Ⅰ　该复合体是电子进入 NAD^+ 呼吸链的门户，由于它在体外能够催化 NADH 的氧化和 CoQ 的还原，因此又被命名为 NADH–CoQ 氧化还原酶。它的主要成分为 NADH 脱氢酶，并以 NADH 为底物，FMN 和铁硫蛋白为辅基。

Quiz5 试设计一个实验，确定一对电子经过 4 种复合物，各能产生多少个质子梯度？

电子在复合体Ⅰ内的流动方向是：NADH → FMN →铁硫蛋白→ CoQ，其中 CoQ 为电子的最终受体。1 对电子在流过复合体Ⅰ时，有 4 个质子泵出线粒体基质进入膜间隙（图 20–7）。

2. 复合体Ⅱ　该复合体是电子从琥珀酸进入 FAD 呼吸链的入口，它在体外能够催化琥珀酸的氧化和 CoQ 的还原，因此又被命名为琥珀酸 –CoQ 氧化还原酶。该复合体最重要的成分是琥珀酸脱氢酶，它也参与三羧酸循环。除此以外，复合体Ⅱ还含有铁硫蛋白和细胞色素 b_{560}。

电子在复合体Ⅱ内的流动方向是：琥珀酸→ FAD →铁硫蛋白→细胞色素 b_{560} → CoQ。CoQ 仍然为电子的最终受体。电子在流过复合体Ⅱ时，无质子离开线粒体基质。

除了琥珀酸脱氢酶能够产生 $FADH_2$ 以外，甘油磷酸脱氢酶和脂酰 –CoA 脱氢酶也能产生 $FADH_2$。但这两种酶产生的 $FADH_2$ 并非是通过复合体Ⅱ进入呼吸链的，而是通过与线粒体内膜结合的其他一些蛋白质将电子直接交给 CoQ（参看脂肪酸氧化），再通过复合体Ⅲ和Ⅳ最终交给氧气。

Quiz6 丙二酸能否抑制呼吸链？为什么？

图 20–7　电子在复合体Ⅰ上的传递

3. 复合体Ⅲ

该复合体的电子供体为 $CoQH_2$。由于它在体外能够催化 $CoQH_2$ 的氧化和细胞色素 c 的还原，因此又名为 CoQ– 细胞色素 c 氧化还原酶。CoQ 作为流动的电子载体，负责从复合体Ⅰ或Ⅱ获取电子，然后“流动”到复合体Ⅲ，并“交出”电子。

电子在复合体Ⅲ内的流动方向是：$CoQH_2$ →细胞色素 b →铁硫蛋白→细胞色素 c_1 →细胞色素 c。细胞色素 c 为电子的最终受体。1 对电子在流过复合体Ⅲ时，通过 Q 循环，有 4 个质子从线粒体基质进入膜间隙。

4. 复合体Ⅳ　该复合体的电子供体为还原型的细胞色素 c。由于它在体外能够催化细胞色素 c 的氧化和 O_2 的还原，因此又被命名为细胞色素 c 氧化酶（cytochrome c oxidase）。

哺乳动物的复合体Ⅳ是由 13 个亚基组成，含有血红素 a、a_3 和两个铜中心——2 个 Cu_A 和 1 个 Cu_B；细菌的复合体Ⅳ则比较简单，只由 3~4 个亚基组成，但在功能上毫不逊色。哺乳动物复合体 Ⅳ中 3 个最大的亚基与细菌的 3 个亚基是同源的，这 3 个最大的亚基由线粒体基因组编码，其他亚基由细胞核基因组编码，起调节作用以及参与复合体的组装。

根据已获得的细菌和牛心肌的细胞色素氧化酶的晶体结构可以看出，亚基Ⅰ是复合体Ⅳ中最大的多肽，含有 12 个跨膜螺旋，但没有任何膜外的结构域。血红素辅基 a 和 a_3 与亚基Ⅰ结合。两个血红素平面与膜垂直。此外，Cu_B 和血红素 a_3 在亚基Ⅰ上形成双核中心（bi-nuclear center），参与电子从血红素 a 到 O_2 的传递；两个 Cu_A 在亚基Ⅱ上与 Cys 残基结合。亚基Ⅱ含有一个大的亲水性结构域突出在膜间隙一侧，参与和还原型的细胞色素 c 的结合。亚基Ⅲ含有 7 个跨膜螺旋，但无任何电子载体，其功能不详。

电子在复合体Ⅳ内的流动方向是：细胞色素 c → Cu_A（亚基Ⅱ）→血红素 a（亚基Ⅰ）→ Cu_B–a_3 双核中心（亚基Ⅰ）→ O_2。一开始，2 个电子传给与双核中心紧密结合的 O_2，形成 O_2^{2-}。随后，另外 2 个电子也传到双核中心，并伴随着 4 个质子被吸收而导致水的形成。O_2 既是复合体Ⅳ的电子最终受体，也是整个呼吸链的电子最终受体。1 对电子在流过复合体Ⅳ时共有 2 个质子泵出线粒体基质进入膜间隙（图 20–8）。已有证据表明，亚基Ⅰ上的带电荷氨基酸残基构成通道，让质子通过它从基质一侧进入膜间隙。

在漫长的生物进化过程中，复合体Ⅳ已形成一种十分有效的机制，可防止氧气部分还原的中间产物如超氧阴离子（$O_2^{\cdot-}$）、过氧化基团（O_2^{2-}）的提前释放。事实上，这些部分还原的产物与双核中心紧密结合，

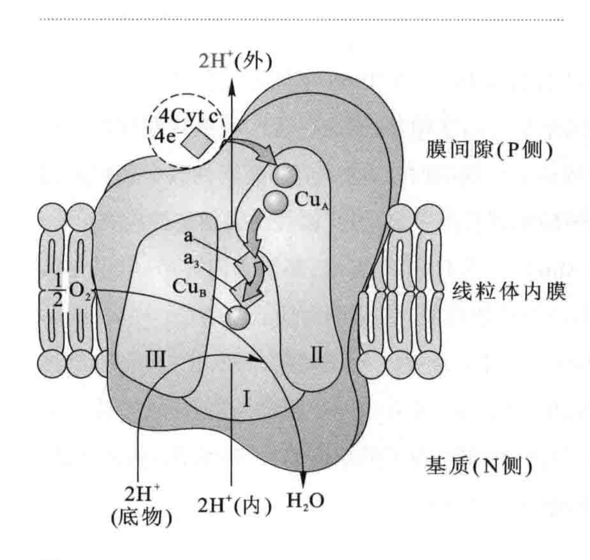


图20-8 电子在复合体Ⅳ上的传递(Nelson 等,2008)

一般不会提前释放,直到水分子的形成。如果没有这种机制,释放出来的 O_2^- 和 O_2^{2-} 就会对细胞造成极大的危害,因为 O_2^{2-} 与 2 个 H^+ 生成 H_2O_2,而 O_2^- 与 H_2O_2 作用生成羟基自由基(·OH)。有很多证据表明,细胞衰老、癌变、DNA 突变和一些疾病的发生均与自由基有一定的关系。

然而,复合体Ⅳ的上述机制并非"完美无缺",有时它也会释放出氧气的部分还原产物,好在细胞内还存在一种由超氧化物歧化酶(superoxide dismutase,SOD)和过氧化氢酶共同构成的"双保险"机制,可及时清除 O_2^- 和 H_2O_2。

$$2O_2^- + 2H^+ \xrightarrow[\downarrow O_2]{SOD} H_2O_2 \xrightarrow{\text{过氧化氢酶}} \frac{1}{2}O_2 + H_2O$$

有趣的是,人体内的某些白细胞在特定的情况下,却可通过呼吸链的"漏洞"产生较多的 O_2^- 和 H_2O_2 用以杀死被吞噬的微生物。

四、植物细胞替代的复合体

许多植物细胞的线粒体除了含有上述 4 种标准的复合体以外,还含有至少两种替代的复合体,一种是复合体Ⅰ的"替补",另外一种是复合体Ⅳ的"替补"。但与"正选"的标准复合体不同的是,电子经过它们进行传递的时候,并不产生跨膜的质子梯度。因此,来自 NADH 中的高能电子的能量更多的是变成了热。其中复合体Ⅰ的"替补"对鱼藤酮不敏感,只有在细胞内[NADH]/[NAD^+]特别高的时候才被激活。复合体Ⅳ的"替补"则对 KCN 不敏感,这也赋予了某些植物具有抗氰呼吸的性质。此外,这种复合体没有铜,但具有双铁氧化中心,在细胞内有大量丙酮酸的时候被激活,直接催化电子从辅酶 Q 传递给氧气。有一种叫巫毒百合(voodoo lily)的植物,其肉穗花序(spadix)含有大量的"替补"复合体Ⅳ,用于局部产热,这有利于花序上挥发性的胺类化合物的释放,产生像腐尸或粪便的臭味,以吸引远处的昆虫前来传粉。

第二节 氧化磷酸化

呼吸链的主要功能是通过与氧化磷酸化的偶联产生能量货币 ATP。当电子沿着呼吸链向下游传递的时候,总伴随着自由能的释放,释放的自由能有很大一部分用来驱动 ATP 的合成,这种与电子传递相偶联的合成 ATP 的方式被称为氧化磷酸化(oxidative phosphorylation,OxP)。

在正常情况下,线粒体内膜上或原核生物质膜上的呼吸链与氧化磷酸化是紧密偶联的。这种偶联有两个方面的含义:一是电子在呼吸链上传递的时候必然发生氧化磷酸化;二是只有发生氧化磷酸化,电子才可能在呼吸链上进行传递。正因为如此,一旦呼吸链被阻断,氧化磷酸化就被抑制,同样,氧化磷酸化被抑制,电子也不可能在呼吸链上正常地传递。

一、氧化磷酸化的偶联机制

有 3 种假说曾被用来解释氧化磷酸化的偶联机制,但只有化学渗透学说被证明是正确的。

化学偶联假说(chemical coupling hypothesis)由 Edward Slater 于 1953 年提出,完全基于底物水平磷酸化的机制,但问题是电子在呼吸链上流动的时候,并没有生成想象中的高能中间物,因此该假说

早已被彻底否定。

构象偶联假说(conformational coupling hypothesis)由 Paul Boyer 于 1965 年提出,它的核心内容是:电子在呼吸链传递的过程中导致某些蛋白质从低能构象变为高能构象,当这些蛋白质从高能构象恢复到低能构象的时候,ATP 就被合成了。虽然电子在呼吸链上传递的时候,的确能够观测到其中某些蛋白质的构象发生变化,但并没有任何证据表明这与氧化磷酸化偶联在一起,因此该假说也很快被抛弃。

化学渗透学说(chemiosmotic hypothesis)由 Peter Mitchell 于 1961 年提出,其核心内容是:电子在沿着呼吸链向下游传递的时候,释放的自由能先转化为跨线粒体内膜(或跨细菌质膜)的质子梯度,随后质子梯度中蕴藏的电化学势能被直接用来驱动 ATP 的合成(图 20-9)。质子多的一侧,正电荷多,故称为 P 侧(positive side),质子少的一侧,负电荷多,故称为 N 侧(negative side)。驱动 ATP 合成的质子梯度通常称为质子驱动力(proton motive force,PMF),是由化学势能(质子的浓度差)和电势能(内负外正)两部分组成,即 $\Delta G=RT\ln(c_2/c_1)+ZF\Delta\psi=2.3RT\Delta\text{pH}+F\Delta\psi$(图 20-10)。

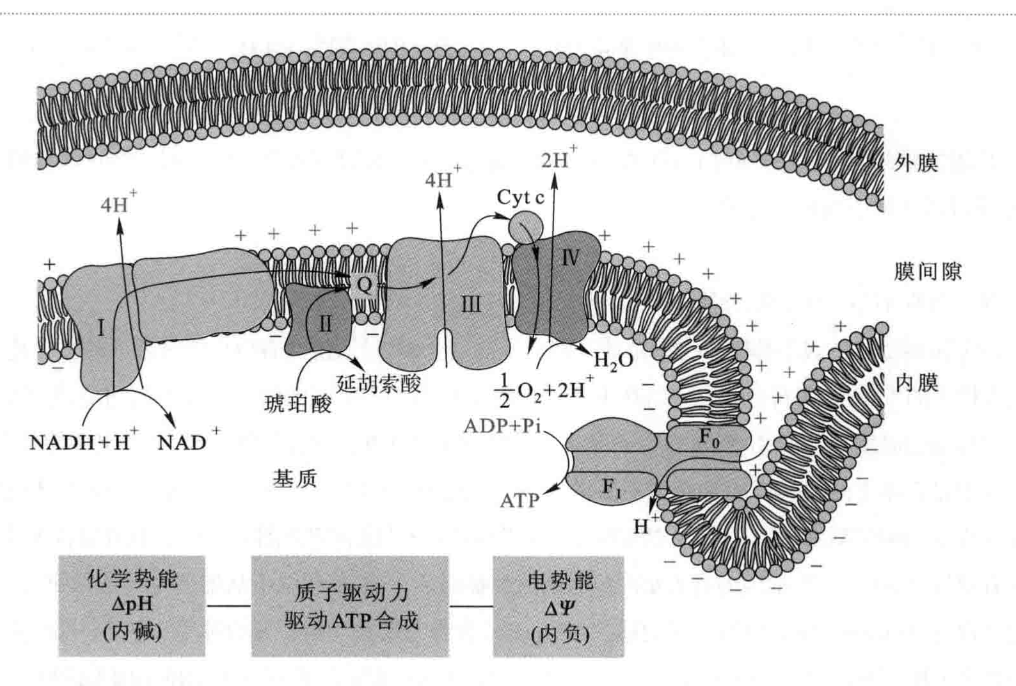


图 20-9 化学渗透学说图解(Nelson 等,2008)

质子驱动力并不总是驱动 ATP 的合成,有时还可以用来直接驱动细胞内物质的主动转运和其他耗能的活动,比如乳糖从胞外进入大肠杆菌、细菌鞭毛的运动以及真核细胞内由核基因编码的蛋白质从细胞质基质转运到线粒体(参看第四十章“蛋白质的翻译后加工及其定向和分拣”)。

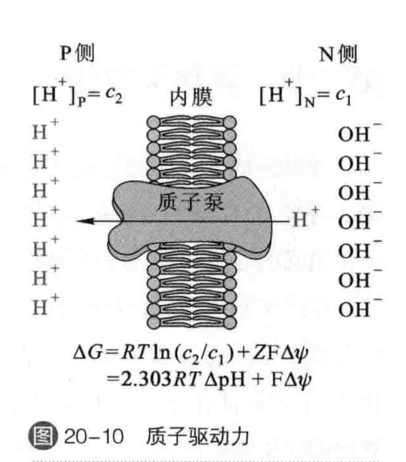


图 20-10 质子驱动力

已有大量证据证明了化学渗透学说的正确性,主要的证据包括:

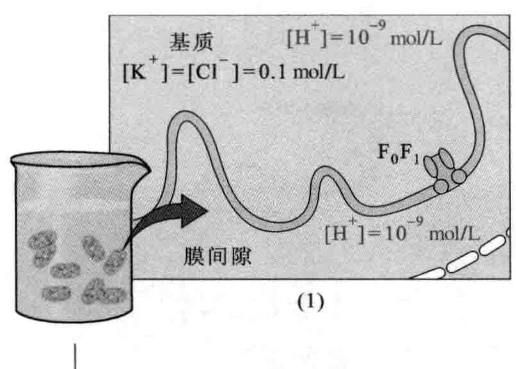
(1) 氧化磷酸化的进行需要完整的线粒体内膜的存在。

(2) 使用精确的 pH 计可以检测到跨线粒体内膜的质子梯度的存在。据测定,一个呼吸活跃的线粒体的膜间隙的 pH 要比其基质的 pH 低 0.75 个单位,内膜两侧的电位差为 0.15~0.2 V。

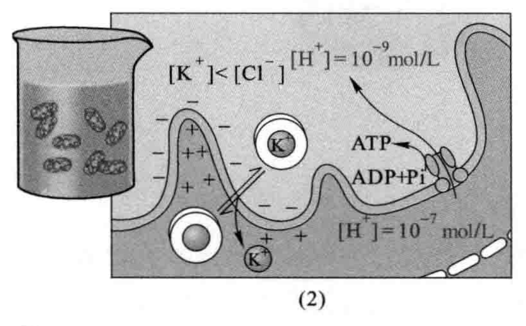
(3) 破坏质子驱动力的化学试剂能够抑制 ATP 的合成。例如缬氨霉素(valinomycin)作为 K^+ 的载体,能够将细胞质基质的 K^+ 带入线粒体基质,而抵消质子驱动力中的电势能,从而抑制 ATP 的合成。

(4) 从线粒体内膜纯化得到的一种酶,能够直接利用质子梯度合成 ATP(详见下一节),此酶称为 F_1F_0-ATP 合酶。

(5) 人工建立的跨线粒体内膜的质子梯度也可驱动 ATP 的合成(图 20-11)。



线粒体悬液pH从9降到7；
有缬氨霉素；无K^+



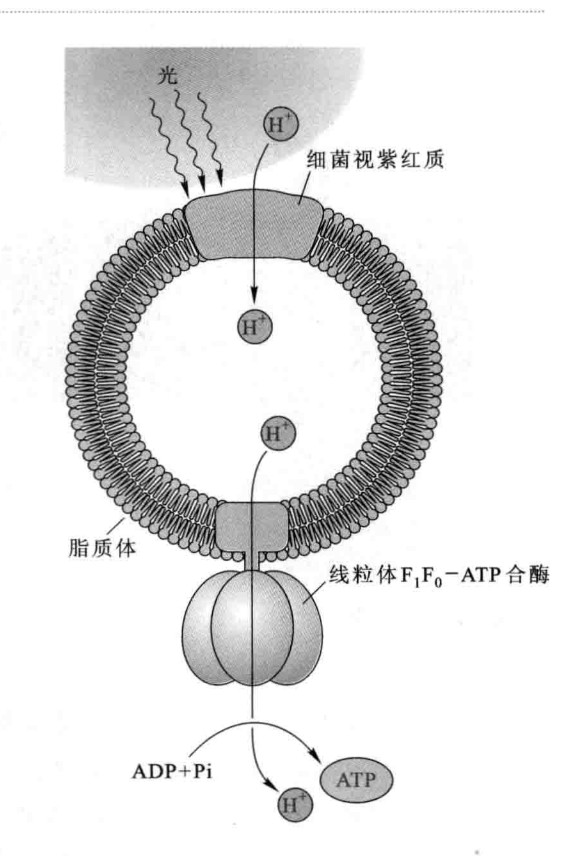


图 20–11 人造质子梯度驱动 ATP 合成的实验（Nelson 等，2008）

不管是通过改变 pH 建立起来的质子梯度，还是通过细菌视紫红质建立起来的质子梯度，都可以被膜上的 F_1F_0–ATP 合酶合成 ATP。

Quiz7 你如何替大肠杆菌制造出质子梯度，为大肠杆菌制造 ATP？

由于获得大量实验证据的支持，化学渗透学说已被人们完全接受，但是，质子梯度究竟是怎样建立的呢？ F_1F_0–ATP 合酶又是如何利用质子梯度合成 ATP 的呢？这两个问题 Mitchell 当时并没有给出很好的解释。现在对于第一个问题，人们一般用“Q 循环”解释复合体Ⅲ产生质子梯度的机制，而复合体Ⅰ产生质子梯度的机制在 2010 年才有突破（参见框 20–1）。至于第二个问题，直到人们对 F_1F_0–ATP 合酶的结构作了充分的研究之后才取得了突破。

Q 循环发生在复合体Ⅲ上，由两个半循环组成。如图 20–12 所示，在第一个半循环中，1 个流动到内膜 P 侧的 QH_2 将其中的 1 个电子经铁硫蛋白和细胞色素 c_1 交给细胞色素 c，同时将 2 个质子释放到膜间隙，第二个电子经细胞色素 b_L 和 b_H 被交回到 N 侧的氧化型 Q 而生成 $Q^{\bar{\cdot}}$，反应式为：

$$QH_2 + \text{Cyt c}(Fe^{3+}) \longrightarrow Q^{\bar{\cdot}} + \text{Cyt c}(Fe^{2+}) + 2H^+(p)$$

紧接着在第二个半循环中，另外一个 QH_2 以同样的方式将 1 个电子经铁硫蛋白和细胞色素 c_1 交给细胞色素 c，同时将 2 个质子释放到膜间隙，第二个电子经细胞色素 b_L 和 b_H 被交给第一个半循环中生成的 $Q^{\bar{\cdot}}$ 并结合基质内的 2 个质子，再生成还原型的 QH_2，并进入下一轮循环，反应式为：

$$QH_2 + Q^{\bar{\cdot}} + 2H^+(n) + \text{Cyt c}(Fe^{3+}) \longrightarrow QH_2 + Q + 2H^+(p) + \text{Cyt c}(Fe^{2+})$$

即：
$$Q^{\bar{\cdot}} + 2H^+(n) + \text{Cyt c}(Fe^{3+}) \longrightarrow Q + 2H^+(p) + \text{Cyt c}(Fe^{2+})$$

总反应是：
$$QH_2 + 2\text{Cyt c}(Fe^{3+}) + 2H^+(n) \longrightarrow Q + 2\text{Cyt c}(Fe^{2+}) + 4H^+(p)$$

由此可以看出，一对电子在复合体Ⅲ经过 Q 循环共有 4 个质子进入膜间隙。

除了复合体Ⅲ能够通过 Q 循环产生跨膜的质子梯度以外，复合体Ⅰ和Ⅳ也能产生跨膜的质子梯度。据测定，1 对电子经过复合体Ⅰ和Ⅳ分别产生 4 个 H^+ 和 2 个 H^+ 梯度，因此，1 对电子经过 NAD^+ 呼吸链或 FAD 呼吸链传给 O_2，可分别产生 10 个 H^+ 和 6 个 H^+ 梯度，按照 4 个 H^+（3 个 H^+ 直接由 F_1F_0–ATP 合酶消耗，1 个 H^+ 在 ATP 与 ADP 的交换运输过程中因为电荷的不平衡而被消耗）从膜间隙返回基质产生 1 分子 ATP 计算，两个呼吸链可分别产生 2.5 个 ATP 和 1.5 个 ATP。

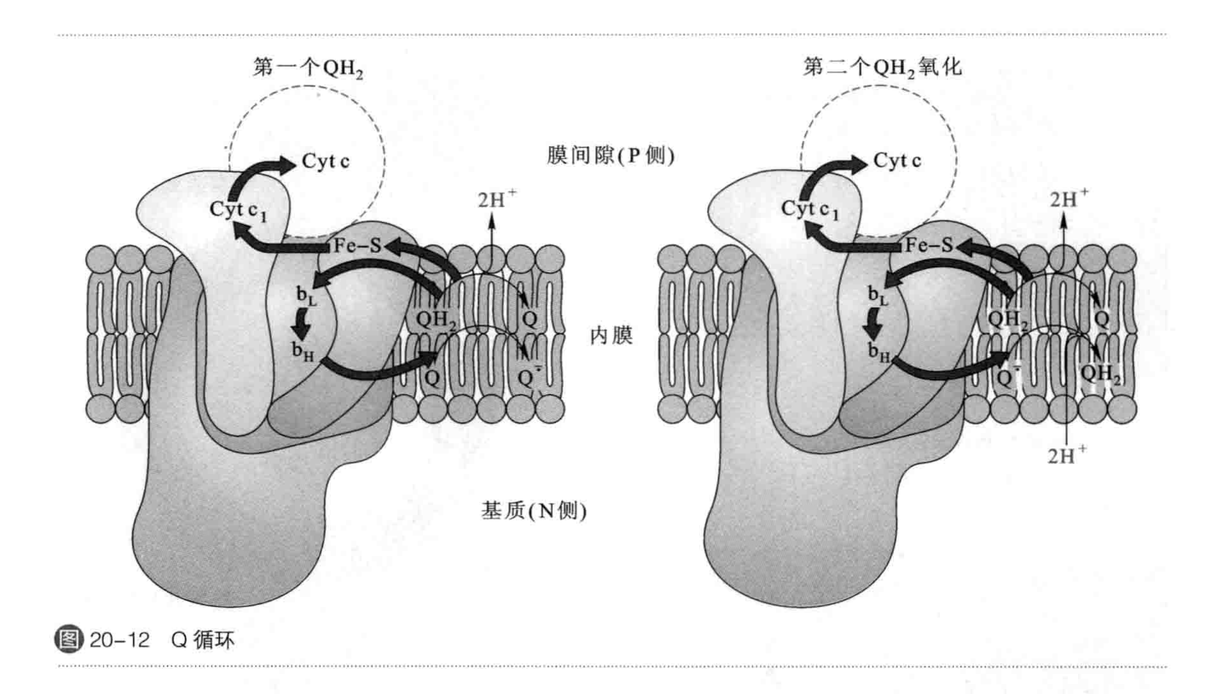


图 20–12　Q 循环

框 20–1　生化研究动态——复合体 I 的结构:细胞中的"蒸汽机"?

不论是真核生物位于线粒体的呼吸链,还是原核生物位于细胞膜上的呼吸链,其中的复合体 I 都是 4 个复合体中最大的,它为氧化磷酸化提供了约 40% 的质子驱动力。已发现,复合体 I 的功能障碍与人类的多种神经退化性疾病有关联。与线粒体上的复合体 I 有 45 个亚基相比,原核生物的复合体 I 要简单得多,只有 14 个核心亚基组成,因此更适合作为研究的对象。

2010 年 5 月 27 日,Efremov R. G. 等人在 *Nature* 上发表了一篇题为"呼吸链复合体 I 的结构:细胞中的"蒸汽机"?(Architecture of Respiratory Complex I:the "Steam Engine" of the Cell)的论文,第一次很好地解释了复合体 I 在将电子从 NADH 传到 CoQ 的过程中如何产生质子梯度的。他们得到了大肠杆菌复合体 I 的跨膜区以及来自一种嗜热菌的完整复合体 I 的晶体结构。通过对晶体结构的分析,他们认为,形如 L 字母的复合体 I 在传递电子的过程中,在两个主要结构域的界面上发生了显著的构象变化。这种构象变化会驱动一段长 α 螺旋发生一种活塞式的运动,使附近 3 段不连续的跨膜螺旋倾斜,从而改变由它们组成的各个质子通道内可解离基团的性质,进而导致 3 个质子转位(图 20–13)。

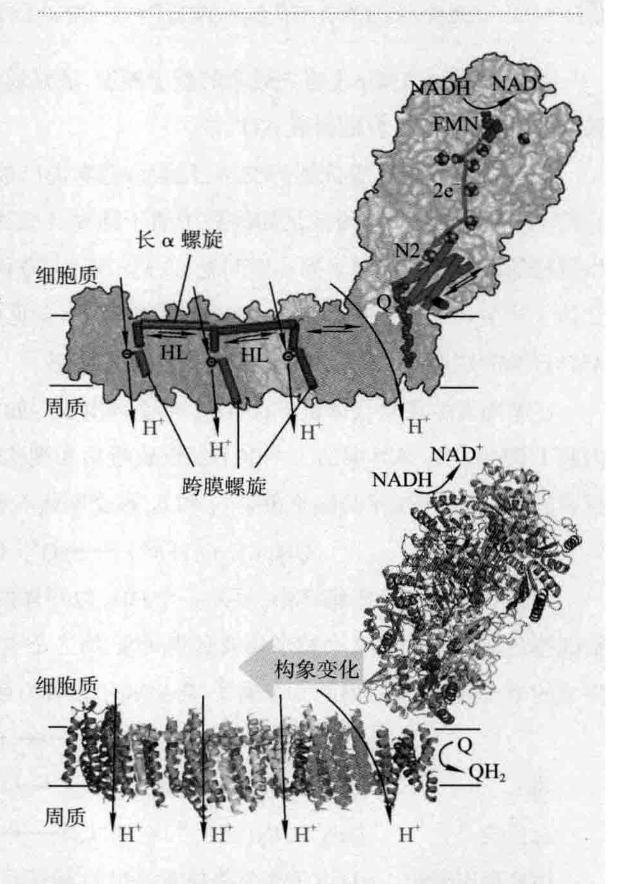


图 20–13　复合体 I 产生质子梯度的分子机制(Nelson 等,2008)

第四个质子可能是在两个结构域之间发生移位的。于是复合体 I 像一个蒸汽机一样,电子传递释放出的能量用来驱动活塞移动,而活塞再驱动一组不连续的螺旋(而不是车轮)移动。

Efremov R. G. 的这篇有关复合体 I 产生质子梯度的论文,被由哈佛大学和剑桥大学等全世界 2 500 名国际顶级医学教授组成的"千名医学家"(Faculty of 1 000 Medicine,F1 000)机构评为当年七大生化论文之首。

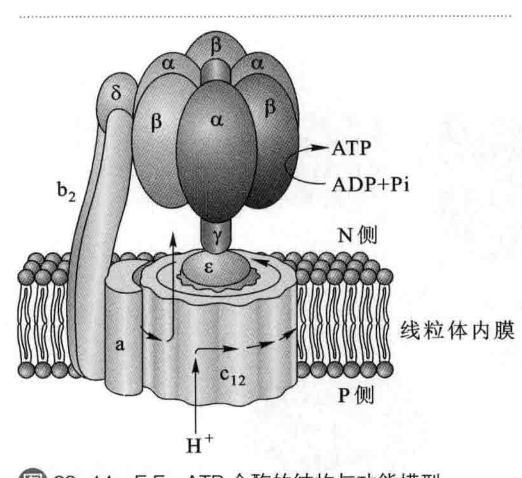


图20-14 F_1F_0-ATP 合酶的结构与功能模型

二、F_1F_0-ATP 合酶的结构与功能

F_1F_0-ATP 合酶与氧化磷酸化直接相关联，因此有必要对它的结构进行较为详细的描述。由于在体外，或者在体内特殊的条件下，它能够催化 ATP 的水解，因此又被称为 F_1F_0-ATP 酶。

F_1F_0-ATP 合酶由 F_1 和 F_0 组成，其中 F_1 含有 3 个 α 亚基、3 个 β 亚基、1 个 γ 亚基、1 个 δ 亚基和 1 个 ε 亚基（可简写为 $\alpha_3\beta_3\gamma\delta\varepsilon$）（图 20-14）。α 亚基和 β 亚基交替排列形成一种环形结构，直接与 ATP 的合成相关。γ 亚基形成一个中央柄，δ 亚基和 ε 亚基直接与 F_0 相互作用。F_0 含有 1 个 a 亚基、2 个 b 亚基和 10~14 个 c 亚基（可简写为 ab_2c_{10}）。所有的 c 亚基横跨内膜作为一个单位（C 单位），与 a 亚基一起，构成一种桶状的质子通道。虽然质子通道是连续的，但有两个不对称的部分，其中一个部分直接向膜间隙开放，另一个部分直接向基质开放。寡霉素（oligomycin）和二环己基碳二亚胺（dicyclohexylcarbodiimide，DCCD）能够直接作用于 F_0 而抑制 ATP 的合成。

温和搅动线粒体内膜可导致 F_1 与 F_0 分离，但这种分离不影响呼吸链上的电子传递，只会使质子梯度不能建立。

F_1F_0-ATP 合酶的主要功能显然是利用质子驱动力来催化 ATP 的合成。但它在体内特殊的条件下，并不是利用质子驱动力来合成 ATP，而是反过来水解 ATP，并利用 ATP 水解释放的能量建立跨膜的质子驱动力。细胞通过这种方式建立质子驱动力，然后再利用质子驱动力，去驱动胞内某些必须直接由电化学势能驱动的耗能活动。像许多微生物在进行无氧发酵的时候，呼吸链已不能运转，但体内的 F_1F_0-ATP 合酶还是有很多，很明显这时候的 F_1F_0-ATP 合酶已“蜕变”成了 F_1F_0-ATP 酶，通过水解由底物水平磷酸化合成的 ATP，为细胞建立所需的跨膜电化学势能。

三、F_1F_0-ATP 合酶的催化机制

1977 年，Paul D. Boyer 提出了“结合变化”学说（binding change hypothesis），该学说能正确地解释 F_1F_0-ATP 合酶的作用机制，并得到了几个关键实验证据的支持。其主要内容是（图 20-15）：

（1）在活性中心合成 ATP 并不需要质子驱动力，与活性中心结合的 ATP 或 ADP 处于平衡。这是“结合变化”学说中最新颖的部分。

人们在对 F_1F_0-ATP 合酶的作用机理真正有所了解之前，很自然地认为该酶在合成 ATP 的时候，肯定是需要消耗能量的，即需要质子梯度的驱动，因为 ATP 是一个高能分子，它的水解即 ATP+H_2O → ADP+Pi 反应的 $\Delta G^{\ominus\prime}$ =-30.5 kJ/mol，这是一个非常大的负值。然而，这个大的负值需要反应发生在标准的条件下，即 ATP、ADP 和 Pi 的浓度均是 1 mol/L，水分子的浓度是 55 mol/L。事实上，如果水分子的浓度只有 10^{-4} mol/L 的话，ΔG=2.2 kJ/mol。这说明在水分子浓度很低的情况下，反应反而有利于 ATP 的合成，更意味着在缺乏水分子的时候，ATP 的合成是不需要能量的。在一个酶的活性中心，水分子可以完全排除在外。F_1F_0-ATP 合酶就是这样的酶，其活性中心主要由疏水氨基酸组成，没有水分子，因此 ADP 和 Pi 可以紧密地结合，并在不需要能量的情况下，自发地合成 ATP（图 20-16）。

（2）合成好的 ATP 离开活性中心却是需要能量的。因此，如果没有质子流过 F_0，与活性中心结合的 ATP 就不会与酶解离。

（3）3 个 β 亚基（β_1、β_2 和 β_3）与 γ 亚基的不同表面结合，不同的 β 亚基将采取不同的构象。在某一时刻，1 个 β 亚基为 T 态，1 个 β 亚基为 L 态，1 个 β 亚基为 O 态。

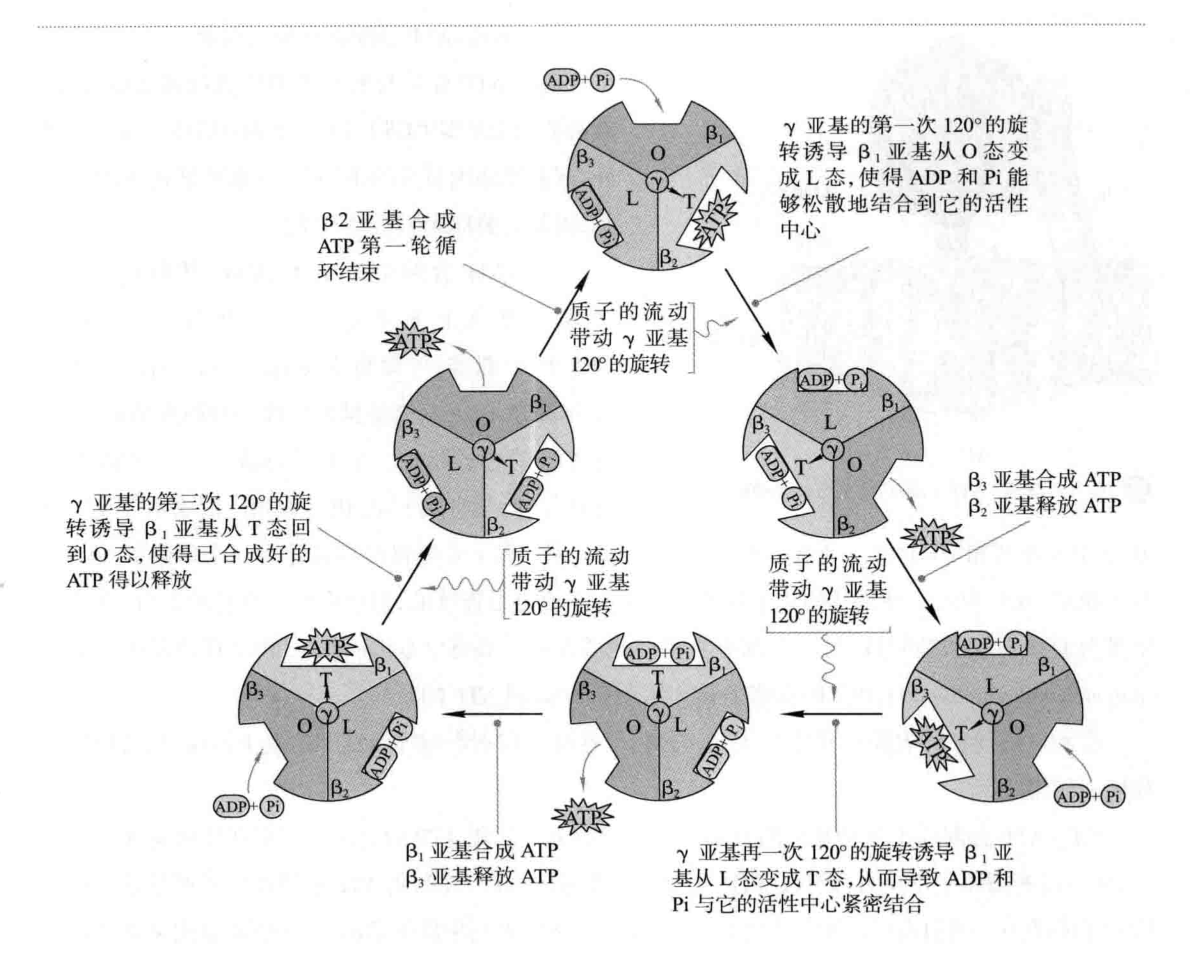


图 20-15 “结合变化学说”图解（Hardin，2010）

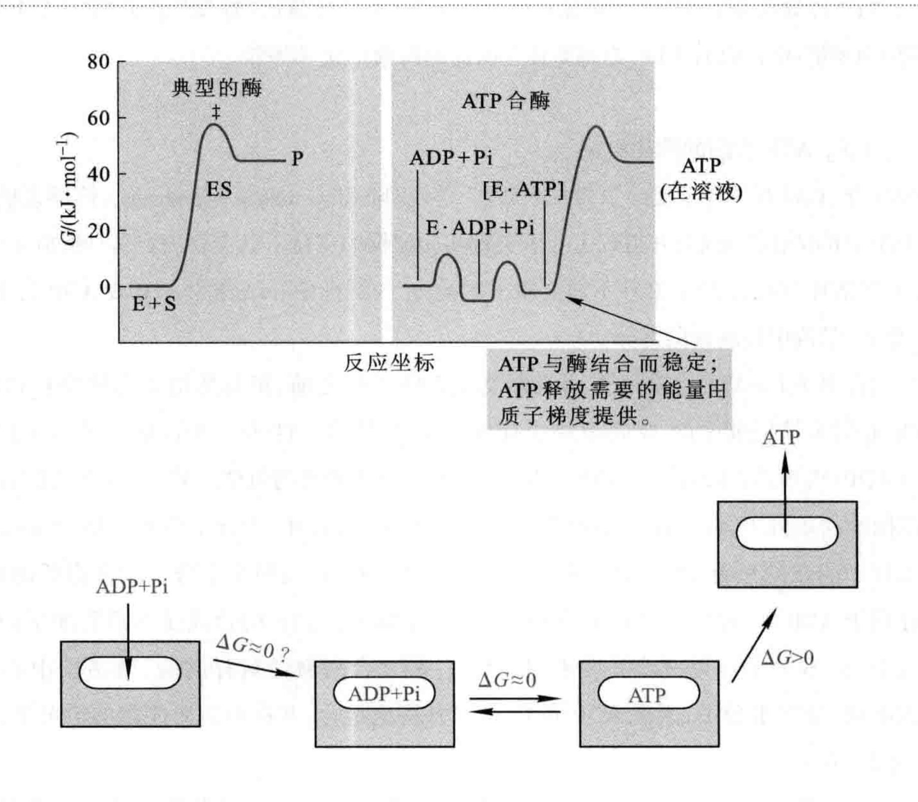


图 20-16 F_1F_0-ATP 合酶催化的 ATP 合成反应的历程

(4) 处于 T 态(tight conformation)的 β 亚基紧密结合 1 分子 ATP,ATP 与 ADP+Pi 处于平衡,但ATP 并不能与它解离;处于 L 态(loose conformation)的 β 亚基结合 ADP 和 Pi,但并不能释放核苷酸;处于 O 态(open conformation)的 β 亚基能够释放结合的核苷酸。

(5) 三种状态的 β 亚基可以相互转变,转变过程由 γ 亚基的转动所驱动。γ 亚基转动的动力来自于质子通过 F_0 的流动,每消耗 1 个质子,γ 亚基转动 120°。

结合变化学说可简化为:质子流动→驱动 C 单位转动→带动 γ 亚基转动→诱导 β 亚基构象变化→ ATP 释放和重新合成。

至于质子是如何通过 F_0 的,有一种观点认为,质子是通过与各 c 亚基上的 Asp 残基可逆结合和解离而进行的(图 20-17)。按照目前的模型,每一个 c 亚基含有一个延伸到质子通道的 Asp 的侧链羧基,当其从膜间隙接受一个质子以后,c 亚基转入疏水的脂膜。转动还将另外一个含有质子的 c 亚基直接暴露到向基质开放的通道内。因为基质有较低的质子浓度,所以 Asp 的侧链羧基将结合的质子释放到基质一侧。转动是受 c 亚基和 α 亚基上一对带相反电荷的基团之间的静电吸引驱动的,其中带负电荷的是 c 亚基上一个 Asp 残基,带正电荷的是 α 亚基上一个 Arg 残基。

该学说已被很多实验所证明,主要的实验证据有:

(1) 图 20-18 所示的同位素交换实验(isotopic exchange experiment)表明,当化学剂量的 ATP、ADP 和无机磷酸与纯化的 F_1 保温在一起时,酶 -ADP/Pi 与酶 -ATP 的交换反应在没有质子梯度的情况下很容易在酶表面发生,$\Delta G^{\ominus\prime}$ 接近 0。这说明与 F_1 紧密结合的 ATP 的合成并不需要质子驱动力,而质子驱动力的作用仅仅在于促进 ATP 的释放。

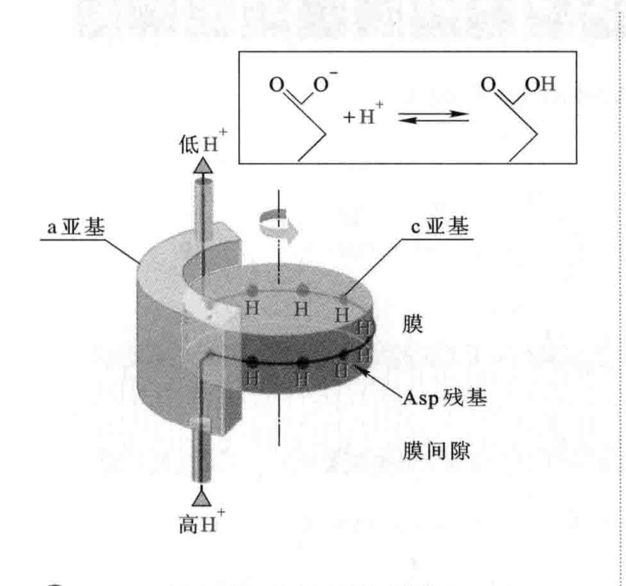


图 20-17　质子通过 F_0 通道回到基质的模型

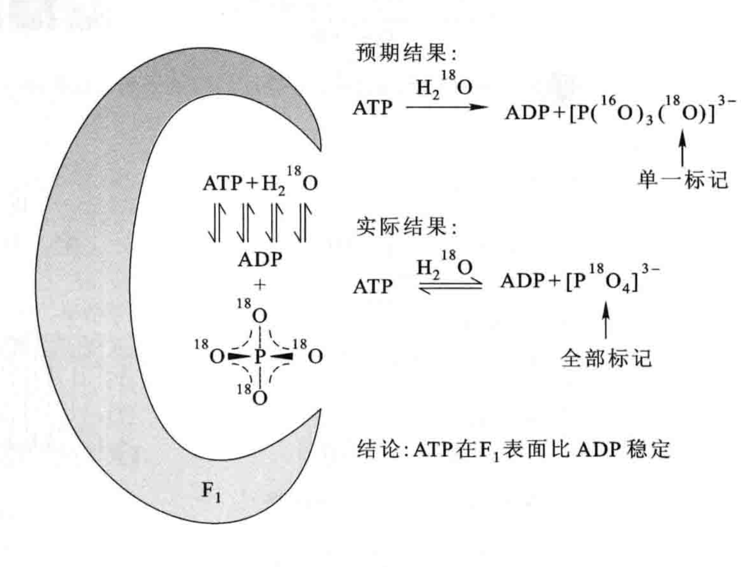


图 20-18　同位素(^{18}O)交换实验(Nelson 等,2008)

(2) 1994 年,由 John Walker 获得的 F_1 的晶体结构清楚地表明,3 个 β 亚基处于不同的构象并和不同的核苷酸配体结合。其中,与 ATP 的非水解类似物(AMPPNP)结合的 β 亚基处于 T 态,与 ADP 结合的 β 亚基处于 L 态,无配体结合的 β 亚基处于 O 态。

Walker 的发现与 Boyer 的结合变化机制是一致的,并为结合变化学说提供了一个有力的证据。Walker 也因此和 Boyer 以及发现 Na^+/ K^+ -ATP 酶的 Jens C. Skou 一起分享了 1997 年的诺贝尔化学奖。

(3) 1997 年,日本的一个研究小组用基因工程的手段对一种嗜热菌的 F_1-ATP 合酶(没有 F_0)进行定向改造,并将其在大肠杆菌细胞里进行表达。他们将 α 亚基和 β 亚基的 N 端添加了组氨酸残基标签(His tag),将 γ 亚基原来的 Cys193 替换成 Ser193,同时将柄状区的 Ser107 替换成 Cys107。Cys107 是 F_1-ATP 合酶分子上唯一的 Cys 残基。在对它进行生物素化(biotinylated)处理以后,可以与链霉生物素结合蛋白(streptavidin)以及被荧光标记的生物素化的肌动蛋白丝结合,形成复合物(1 个

链霉生物素结合蛋白有 4 个生物素的结合位点)。再将改造过的酶与预先喷涂金属 Ni 复合物的载玻片保温一定时间后,F_1-ATP 合酶通过它的组氨酸标签与金属 Ni 复合物之间的螯合作用被锚定在载玻片上。然后,加入亲和素(streptavidin)及被荧光标记的生物素化的肌动蛋白丝(图 20-19)。由于生物素与亲和素之间有高度的亲和力,荧光标记的肌动蛋白丝被固定在 γ 亚基的柄部。被荧光标记的肌动蛋白丝比原来的 F_1 要大好多倍,因此能够在光学显微镜下观测到。这时再加入 2 mmol/L ATP,可以在荧光显微镜下,直接观测到肌动蛋白丝沿逆时针方向有规律地转动(图 20-20)。这种旋转与细菌的鞭毛受质子梯度的转动相似。上述实验说明,ATP 的水解在体外可以驱动 F_1F_0-ATP 合酶 γ 亚基的旋转。同时,它也意味着,质子驱动力在体内同样可以驱动 F_1F_0-ATP 合酶 γ 亚基的旋转,导致 ATP 的释放和合成。因此,有人将 F_1F_0-ATP 合酶称为世界上最小的"分子涡轮发电机"或"分子印钞机"。正是该"印钞机"的持续不断地运转,"能量货币"才可能源源不断地被制造出来。

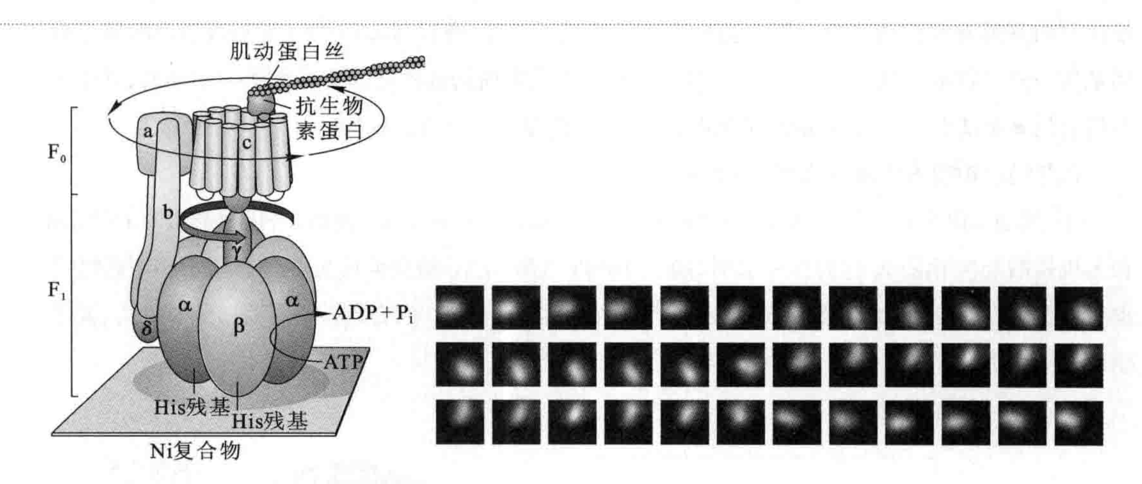


图 20-19 日本科学家的 γ 亚基的旋转实验及其 γ 亚基的旋转轨迹(Nelson 等,2008)

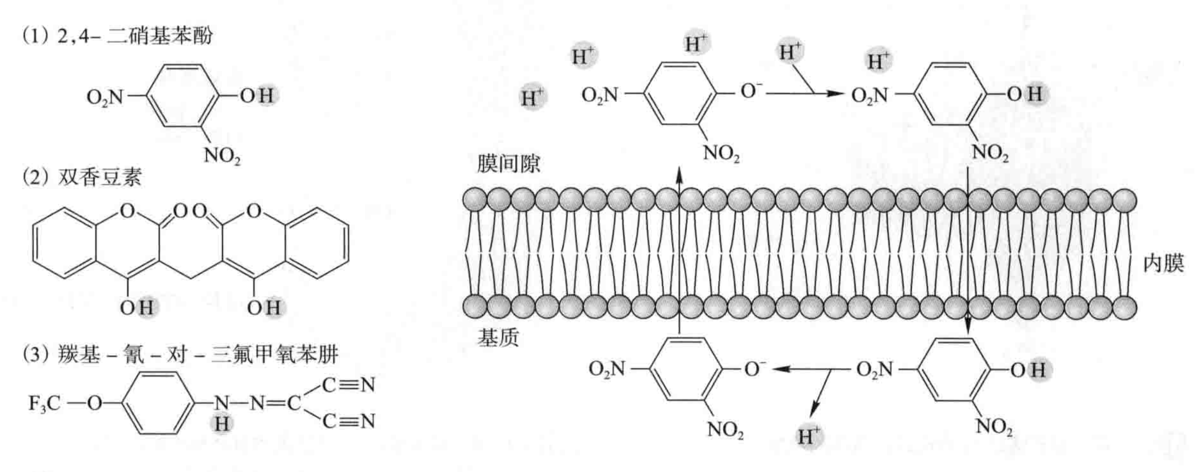


图 20-20 三种解偶联剂的化学结构以及 DNP 的作用机制

四、氧化磷酸化的解偶联

氧化磷酸化与呼吸链通常是紧密偶联的,但低水平的质子泄漏时刻发生在线粒体内膜上。因此,确切地说,线粒体通常是部分解偶联的,而完全的解偶联一般是解偶联剂(uncoupler)的作用所致。

解偶联剂的作用机制在于它们能够快速地消耗跨膜的质子梯度,使得质子难以通过 F_1F_0-ATP 合酶上的质子通道合成 ATP,从而将贮存在质子梯度之中的电化学势能全部变成了热。此外,随着质子梯度的消失,电子在呼吸链上的"回流"压力将会减轻,进而使细胞内脂肪和糖类等物质的生物氧化更加旺盛。也正因为如此,有的解偶联剂曾被用做减肥药,例如 20 世纪 30 年代,2,4- 二硝基苯酚(dinitrophenol,DNP)就曾被人们作为减肥药服用。

有两类解偶联剂：一类为有机小分子化合物，通常为脂溶性的质子载体（proton ionophore），带有可解离的基团（pK_a 约为 7.2），例如 DNP、双香豆素（dicumarol）和羰基－氰－对－三氟甲氧基苯肼（FCCP）等。这些有机小分子在线粒体膜间隙相对低的 pH 环境中，可以接受质子，并主要以非解离的形式存在。自身具有的脂溶性使得它们很容易自由扩散到基质一侧。当它们到达基质一侧以后，较高的 pH 环境促使它们解离并释放出在膜间隙结合的质子。质子梯度就这样没有经过 F_1F_0-ATP 合酶白白地被消耗掉了（图 20-20）。阿司匹林和甲状腺素的结构由于与 DNP 相似，因此在特定情况下也可以作为解偶联剂来发挥作用。另一类为天然的解偶联蛋白（uncoupling protein，UCP）。

UCP 在线粒体内膜上形成质子通道，使质子流发生"短路"，不通过 F_1F_0-ATP 合酶就能返回到线粒体基质。至少已在高等动物体内发现 5 种类型的 UCP（UCP1~UCP5）。UCP1 又名产热素（thermogenin），它主要存在于动物的褐色脂肪组织（brown adipose tissue）（褐色源于高含量的线粒体），与机体的非颤抖性产热有关（nonshivering thermogenesis）。机体对寒冷作出的反应机制是：交感神经末梢释放去甲肾上腺素，去甲肾上腺素再激活褐色脂肪组织中的脂肪酶。脂肪酶水解脂肪，释放出游离的脂肪酸（参看第二十九章"脂肪酸代谢"）。游离脂肪酸不仅可作为代谢燃料经氧化产生 ATP 和质子梯度，还能与 $CoQH_2$ 和嘌呤核苷酸一起直接激活产热素。一旦产热素被激活，质子流发生就会"短路"生热。UCP2 存在于多数细胞中，UCP3 主要存在于骨骼肌，UCP4 和 UCP5 存在于脑，它们都是受到高度调控的蛋白质，一旦被激活，都能增加热量的产出。由于它们会影响到代谢效率，其含量偏低（特别是骨骼肌中的 UCP3）可能是某些个体或人群肥胖的原因。

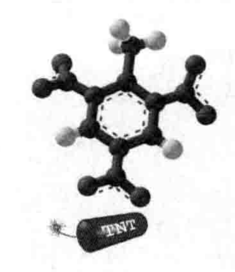
Quiz8 有人认为失去 F_1 的 F_1F_0-ATP 合酶将成为一种解偶联蛋白，你支持这种观点吗？为什么？

UCP 也存在于植物体内，不过通常被称为植物线粒体解偶联蛋白（plant uncoupling mitochondrial protein，PUMP）。PUMP 的功能主要是防止线粒体膜电位过高而形成太多活性氧（ROS）。当线粒体内的代谢物过剩的时候，植物通过构成呼吸链复合物的替补以及 PUMP 的作用，来增加对过剩代谢物的代谢，将多余的能量转变成热。某些植物体面对冷胁迫时会产生大量的 PUMP。

框 20-2 生化趣事——炸药与减肥

第一次世界大战期间，三硝基甲苯（TNT）作为炸药已在战争中大量使用，在炮弹和炸弹里面装载的都是它。奇怪的是，当时有人注意到，许多在军工厂里负责将 TNT 装载到炮弹和炸弹里面的妇女身体消瘦，而且经常发热，对此原因一直不明。在很长一段时间以后，人们发现了一个与 TNT 非常类似的化学品，即 DNP，它在体内作为氧化磷酸化的解偶联剂，可造成跨线粒体膜的质子泄漏，也能引起发烧和体重减轻。

显然，从结构上来看，TNT 并不能直接作为解偶联剂，原因是在它的苯环上没有亲水的可解离羟基。然而，一旦它被吸入到体内，在细胞内经内质网膜上的细胞色素 P450 的解毒处理，即羟基化修饰，便变成了与 DNP 类似的解偶联剂。



五、氧化磷酸化的抑制

某些化学试剂能够直接或间接抑制细胞内的氧化磷酸化，这些抑制剂分为 4 类（表 20-3）：①呼吸链位点特异性抑制剂。通过抑制呼吸链上的电子传递，阻止质子梯度的生成而间接抑制氧化磷酸化，例如抗霉素 A、鱼藤酮和氰化物等；②F_1F_0-ATP 合酶特异性抑制剂。直接作用 F_1F_0-ATP 合酶而导致 ATP 不能被合成，如寡霉素和 DCCD 等；③解偶联剂。破坏 pH 梯度，使 ATP 不能合成；④ATP/ADP 交换体或转位酶的特异性抑制剂。通过抑制线粒体内外 ATP/ADP 的交换而间接抑制氧化磷酸化，如苍术苷（atractyloside）和米酵菌酸（bongkrekik acid）。

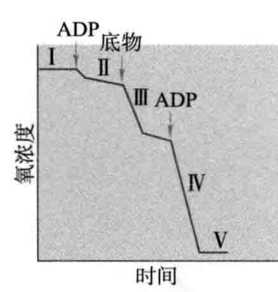
表 20-3 氧化磷酸化的抑制剂

抑制类型	抑制剂名称	作用位点或作用机制
呼吸链抑制剂	鱼藤酮、安米妥、杀粉菌素	复合体Ⅰ
	萎锈灵	复合体Ⅱ
	抗霉素 A	复合体Ⅲ
	氰化物、CO、H_2S、叠氮化物	复合体Ⅳ
F_1F_0-ATP 合酶抑制剂	金轮霉素(aurovertin)	抑制 F_1
	寡霉素、杀黑星菌素(venturicidin)	抑制 F_0
	DCCD	阻止质子通过 F_0 通道
解偶联剂	DNP、FCCP	脂溶性质子载体
	缬氨霉素	钾离子载体,破坏电势能
	生热素	质子通道
ATP/ADP 交换体抑制剂	苍术苷、米酵菌酸	抑制线粒体基质内的 ATP 与细胞质内的 ADP 之间交换

六、P/O 值与氧化磷酸化的调节

氧化磷酸化的效率可以通过测定 P/O 值来确定。P/O 值是指在电子传递过程中,每消耗 1 mol 氧原子所消耗的无机磷酸的物质的量(mol)。消耗的氧原子数目相当于传递给氧气的电子数的 1/2,消耗的无机磷酸等于氧化磷酸化产生的 ATP。因此 P/O 值越高,氧化磷酸化的效率就越高。1 对电子经过 NADH 呼吸链或 $FADH_2$ 呼吸链产生的 ATP 数目是不一样的,两条呼吸链的 P/O 值也就不同。按照前面的计算原则,NADH 呼吸链的 P/O 值为 2.5,$FADH_2$ 呼吸链的 P/O 值为 1.5。

细胞内的氧化磷酸化是受到严格调控的,调控的手段主要是氧化磷酸化对电子传递的反馈。这实际上是由 ADP 浓度控制的,这种由 ADP 对氧化磷酸化的调节被称为呼吸控制(respiratory control)。当线粒体内[ADP]↑,F_1F_0-ATP 合酶活性↑,质子梯度↓,电子传递速率↑,耗氧率↑,细胞呼吸↑;相反,当线粒体内[ADP]↓,F_1F_0-ATP 合酶活性↓,质子梯度↑,电子传递速率↓,耗氧率↓,细胞呼吸↓。



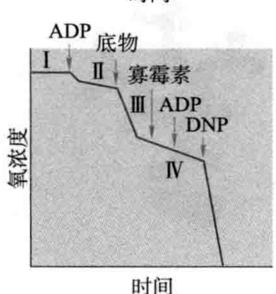


图 20-21 ADP 和底物浓度对呼吸的影响

氧气作为呼吸链的最终电子受体,它的消耗程度能直接反映电子传递和氧化磷酸化的效率。下面的实验说明了[ADP]对细胞呼吸和氧化磷酸化的影响。如图 20-21 所示,可以在密闭环境中,根据离体线粒体在不同条件下的氧耗情况(氧气的消耗情况可通过氧电极直接测出),将细胞呼吸功能分成 5 种状态:

(1) 状态Ⅰ 可氧化底物(如苹果酸)很少,无 ADP,这时耗氧率极低。

(2) 状态Ⅱ 刚刚补充 ADP 以后的状态。由于线粒体内部存在少量的内源底物,因而可以观测到刚加入的 ADP 可短暂刺激氧气的消耗。但随着内源底物的耗尽,电子传递和氧化磷酸化很快停止,因此不会再有氧气的消耗。

(3) 状态Ⅲ 同时加入 ADP 和底物。由于加入的底物可以立即被线粒体氧化利用,释放的电子进入呼吸链,ADP 则被氧化磷酸化利用,因此,这时可以观测到耗氧率的急剧攀升直至加入的 ADP 被耗尽。

(4) 状态Ⅳ 加入的 ADP 已被耗尽,耗氧率维持较低水平。如果这时再次加入 ADP,可以发现耗氧率再次大幅上升。

(5) 状态Ⅴ 氧气完全被耗尽,这时即使还残留底物和 ADP,电子传递和氧化磷酸化也被抑制了。但是,在上面的实验中,如果在状态Ⅳ时加入寡霉素,这时的耗氧率并没有显著的提高。这是因为,寡霉素直接抑制氧化磷酸化,而氧化磷酸化与电子传递是偶联的,当氧化磷酸化被抑制时,电子传递也就被抑制了,这时氧气作为电子的最终受体也就成为多余的了。如果在状态Ⅳ同时加入寡霉素和 DNP,则会发现耗氧率迅速提高,这是因为 DNP 作为解偶联剂可以解除氧化磷酸化与电子传递之间的偶联关系,即使氧化磷酸化被抑制了,电子照样能传递,氧气依旧被消耗。

上述 5 种状态中,只有状态Ⅲ和Ⅳ为正常的生理条件下线粒体经常处于的呼吸状态。其中状态Ⅳ为细胞处于静息条件(resting condition)下的呼吸状态,这时细胞内的能荷较高,[ADP]较低,细胞

对新 ATP 合成的需要有限,因此耗氧率不高。当细胞需要大量 ATP 的时候,例如剧烈的肌肉运动,ATP 的消耗会引起细胞质基质内的[ADP]迅速提高,此时在线粒体内膜上的 ADP/ATP 交换体的作用下,细胞质基质内的 ADP 与线粒体内的 ATP 发生交换,最终导致线粒体内的[ADP]提高,线粒体随即进入状态Ⅲ,耗氧率随之剧增。

科学故事——结合变化学说的发现

ATP 作为细胞通用的能量货币,其合成机制一直是科学家研究的焦点课题之一。Mitchell 因提出氧化磷酸化的"化学渗透"学说而获得 1978 年的诺贝尔化学奖。在他的学说中,虽然正确地预测了电子在呼吸链上传递释放的能量转变成跨膜的质子梯度,但并没有正确解释 F_1F_0-ATP 合酶是如何利用质子梯度合成 ATP 的。

早在 20 世纪 50 年代早期,就有人使用 ^{18}O 发现线粒体能够催化 Pi 中的 O 与水分子上的 O 进行快速交换。当 Boyer 成为明尼苏达大学的一名研究者以后,也开始使用 ^{18}O 和 ^{32}P 研究氧化磷酸化的过程。他发现整个过程是动态可逆的,而且 ^{18}O 交换比 ^{32}P 交换要快。此结果以及其他的一些研究成果促使他和其他一些研究人员,去寻找氧化磷酸化过程中可能存在的高能中间物,以支持氧化磷酸化的"化学偶联"学说,但都以失败告终。

于是,Boyer 将其研究的重心转移到其他的酶,但并没有完全放弃对 ATP 形成的研究。到了 70 年代,Boyer 当上了刚刚成立的 UCLA 跨系科的分子生物学研究所的首任所长。在一次学术报告上,他的思绪重新回到以前一些始终没有得到解释的氧交换数据上。忽然,一种全新的氧化磷酸化的概念闪现在他的脑海里:他寻找的中间物可能并不存在,也许有一种意想不到的机制。他回忆到:"我清醒地意识到,如果来自氧化的能量并不用来制造 ATP 分子,而是用来驱动紧密结合的 ATP 释放,那么结果就能被解释了!"这是他第一次对 ATP 合酶制造 ATP 的机制有了满意的认识。但是,他的新设想并没有很快被接受。Boyer 递交给 *The Journal of Biological Chemistry* 的相关论文被拒就是证据。

怀疑甚嚣尘上。但是,Boyer 和他的研究助手感到他们在正确的轨道上。后来,Boyer 行使了他刚刚当上的美国科学院院士不需要同行审稿就能发表一篇论文的特权,使他的论文终于发表了。他解释到:"感觉到,它就是对的。"于是,这篇论文成为了后来被称为 ATP 合成"结合变化"机制的开始。在随后的十几年里,Boyer 的研究小组发现了 ATP 合酶的另一种不同寻常的特征,即它的 3 个催化亚基总是按照一种有序的方式进行催化。到了 80 年代早期,在这个领域的许多研究者开始接受他提出的学说。但是,另一个由 Boyer 提出的有争议的酶学概念还没有得到证明。^{18}O 交换研究显示,3 个催化位点上的催化行为一样。催化亚基似乎组织成一个圆圈,并受到内部的一个亚基影响。Boyer 提出,里面的一个亚基就像汽车上的凸轮轴(cam shaft),通过转动导致催化位点经历 3 个步骤:ADP 和 Pi 结合→ATP 形成和收紧→ATP 释放。在持续不断的旋转催化过程中,所有的 3 个催化位点在任何时候都处于不同的构象。

几千种酶已被研究过,但还没有人发现哪一种酶能进行旋转催化。Boyer 又开始冒险了!他也承认:"这个想法没有得到多大的热情"。但很快,Boyer 的研究小组得到了支持旋转催化的数据,可是缺乏使它们令人信服的评估技术。到了 80 年代后期,Boyer 已进入古稀之年,他决定让其他人去解决旋转催化的问题。

John Walker 是英国剑桥分子生物学医学研究学会实验室的一位资深科学家。在 1994 年,他和他的同事根据 X 射线晶体分析,确定了 ATP 酶催化亚基的三维结构。Walker 的结构支持了 Boyer 提出的旋转催化的概念。另一个证据来自于日本的 Masasuke Yoshida 及其同事,他们使用特殊的方法在显微镜下观察到亚基的旋转。

Boyer 后来和 Walker 因 ATP 合酶的研究而获得了诺贝尔奖。对于 Boyer,这正是一个美妙的时刻!科学的发现就是这样:先是怀疑新的理论,然后是欣赏它。Boyer 曾说:"我是幸运的,因为我研究的酶不但重要,而且具有新的催化性质"。

e20-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Oxidative_phosphorylation（维基百科有关氧化磷酸化的内容）
2. http://themedicalbiochemistrypage.org/oxidative-phosphorylation.html（完全免费的医学生物化学课程网站，有关氧化磷酸化的内容）
3. http://www.wiley.com/college/boyer/0470003790/animations/electron_transport/electron_transport.htm（内有呼吸链电子传递的动画，形象生动）
4. http://www.chemistry.wustl.edu/~edudev/LabTutorials/Cytochromes/cytochromes.html（美国华盛顿大学圣路易斯分校化学系提供的有关氧化磷酸的内容，浅显易懂）

参考文献：

1. Sazanov L, et al. The architecture of respiratory complex I. *Nature*, 465:441-445, 2010.
2. Yankovskaya V, et al. Architecture of succinate dehydrogenase and reactive oxygen species generation. *Science*, 299:700-706, 2003.
3. Boyer P D. The ATP synthase-a splendid molecular machine. *Annu Rev Biochem*, 66:717-749, 1997.
4. Yasuda R, et al. Direct observation of the rotation of F_1-ATPase. *Nature*, 386:299-302, 1997.
5. Boyer P D. A perspective of the binding change mechanism for ATP synthesis. *FASEB J*, 3(10):2 164-2 178, 1989.
6. Mitchell P. Coupling of phosphorylation to electron and hydrogen transfer by a chemi-osmotic type of mechanism. *Nature*, 191(4 784):144-148, 1961.

第二十一章 生物大分子的消化和吸收

分解代谢的第一步是生物大分子水解成其组成单位。对于动物来说,由食物中摄入的生物大分子的分解代谢始于消化道。通过消化道分泌的各种水解酶的消化作用,糖类、脂质、蛋白质和核酸被降解成各自的组成单位,再经过吸收或转运进入细胞,在细胞内可进一步被分解。

本章主要介绍动物消化道内发生的各类生物大分子水解和吸收的生化机理。

第一节 糖类的消化和吸收

食物中含有的糖类有单糖、蔗糖、乳糖、糖原、纤维素和淀粉等,除了单糖可以直接被小肠上皮细胞吸收以外,其他糖类吸收前都需要先被水解。

一、双糖、寡糖和多糖的酶促降解

(一) 多糖的水解

除了反刍动物和白蚁以外,绝大多数动物的消化道缺乏水解β-1,4 糖苷键的水解酶,因此带有如此连接方式的多糖就无法被水解利用,例如纤维素和几丁质。但纤维素作为膳食纤维的主要成分,具有刺激肠蠕动、清洁消化壁和增强消化的功能,同时还可稀释和加速食物中的致癌物等有毒物质的清除,有利于保护脆弱的消化道和预防结肠癌,因此,它也是维持健康必不可少的物质。几丁质则有纤维素所不具备的特性,它是唯一含阳离子的可食性动物纤维,经常食用也是有益于健康的。

带有α-1,4 糖苷键的淀粉和糖原则可以在唾液或胰液内的α-淀粉酶(α-amylase)的催化下,水解成较为简单的产物,或者在外切-1,4-α-糖苷酶(exo-1,4-α-glucosidase)即葡糖淀粉酶(glucoamylase)的催化下,依次释放出葡萄糖。α-淀粉酶属于内切酶,仅能水解α-1,4 糖苷键,对分支点上的α-1,6 糖苷键无能为力。因此,由α-淀粉酶水解的产物主要是麦芽糖、异麦芽糖、麦芽三糖和α-极限糊精(α-limit dexitrin)(图 21-1)。α-极限糊精约由 8 个葡萄糖单位组成,带有 1~2 个α-1,6 糖苷键,可在小肠黏膜表面的寡聚-1,6-糖苷酶(oligo-1,6-glucosidase)或异麦芽糖酶(isomaltase)的催化下,丢掉以α-1,6-糖苷酶连接的葡萄糖单位。

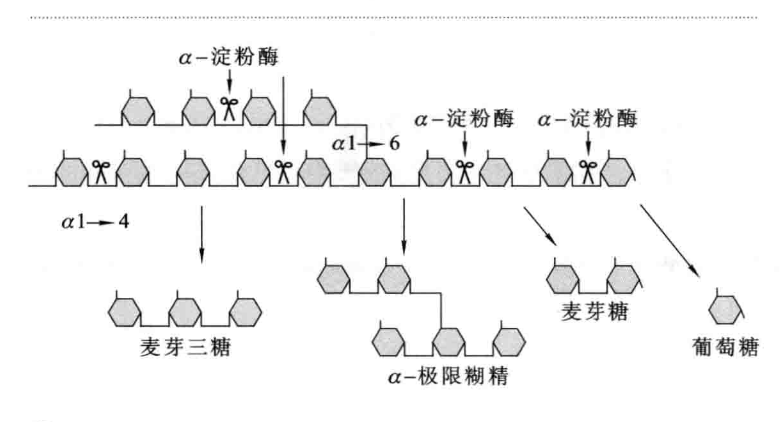


图 21-1 淀粉在消化道内的酶促水解

(二) 双糖和寡糖的水解

从食物中直接摄入的或由多糖降解产生的双糖和寡糖还需要进一步水解。参与这一过程的酶有:①蔗糖-α-糖苷酶(sucrose-α-glucosidase),即蔗糖酶(sucrase),专门将蔗糖水解成葡萄糖和果糖;②β-半乳糖糖苷酶(β-galactosidase),即乳糖酶(lactase),将乳糖水解成半乳糖和葡萄糖;③麦芽糖酶,将麦芽糖和麦芽三糖水解成葡萄糖;④异麦芽糖酶,将异麦芽糖和 α-极限糊精水解成葡萄糖;⑤α,

α- 海藻糖酶(α,α-trehalase),将海藻糖水解成葡萄糖。这些消化酶都固定在肠细胞的刷状缘(brush border)上,在它们的作用下,最后的产物即高浓度的单糖将集中在这一特殊的区域,为吸收提供了便利。

二、单糖的吸收和转运

单糖进入细胞被吸收的过程受载体蛋白即转运蛋白的介导(transporter-mediated),具有底物特异性和立体特异性,遵循饱和动力学,并受到特定抑制剂的抑制(参看第十六章第三节"物质的跨膜转运")。

至少有两个跨膜转运系统参与单糖由肠腔进入小肠上皮细胞内的过程(图 21-2):

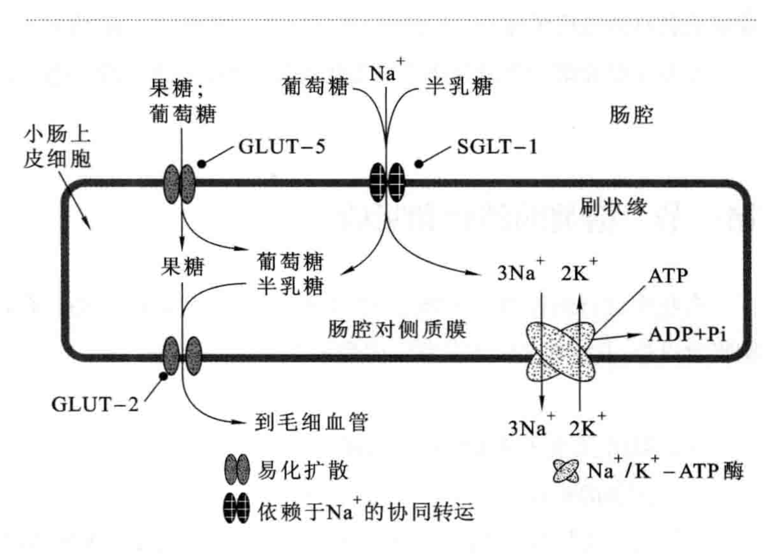


图 21-2 单糖的跨膜转运

(1) 属于次级主动转运的 Na^+- 单糖共转运蛋白(Na^+-monosaccharide transporter,SGLU)系统 该系统的载体蛋白是个四聚体蛋白,每一个单体对 D- 葡萄糖、α- 甲基 -D- 葡萄糖和 D- 半乳糖专一。单糖进入细胞需要 Na^+ 的伴随,实际上,Na^+ 顺着浓度梯度从胞外进入胞内的过程为葡萄糖或半乳糖的吸收提供了动力。进入细胞内的 Na^+ 在 Na^+/K^+ 泵的催化下再离开细胞。

(2) 属于被动转运的且不依赖于 Na^+ 的易化扩散转运系统(GLUT5) 该系统对 D- 果糖专一。通过此系统进入肠细胞的果糖被转变成葡萄糖。此外,在小肠上皮细胞肠腔对侧的质膜上,存在另外一种不依赖于 Na^+ 但能转运 3 种常见的单糖(D- 葡萄糖、D- 半乳糖和 D- 甘露糖)以及 2- 脱氧 -D- 葡萄糖的易化扩散转运系统(GLUT2)。GLUT2 还存在于肝和肾中,而 GLUT 家族的其他成员则存在于所有的细胞中,参与葡萄糖的易化转运。

Quiz1 为什么葡萄糖从消化道进入小肠上皮细胞需要通过主动转运途径,而离开小肠上皮细胞进入血液却是易化扩散?

在正常的生理条件下,位于消化道、肝和肾的 GLUT2 负责将葡萄糖运出细胞,进入血液。而红细胞和脑细胞上的 GLUT1、脂肪组织和肌肉组织的 GLUT4 则主要负责葡萄糖的吸收,即将血液中的葡萄糖转运到各自的细胞内。

如果单糖的吸收被破坏,那大量的单糖分子就会聚集在小肠上皮细胞的表面,产生高渗环境,影响水的吸收,从而引起腹泻。

第二节 脂质的消化和吸收

无论是水解还是吸收,脂质与糖类或蛋白质都有很大的不同。差别产生的原因与脂质本身的脂溶性(水不溶性)有关。

一、脂质的酶促降解

脂质的水不溶性给其水解和吸收都造成了很大的困难。首先,就水解而言,由于参与水解的酶,例如脂肪酶、磷脂酶和胆固醇酶都是水溶性的,因此,脂质的水解只能在水相和油相的界面进行。显

然，油相与水相的接触面积越大，水解的速率就越快。那么，如何能增加消化道内脂质与水相的接触面积呢？其次，就吸收而言，由于水解产物仍然倾向于聚集在一起形成更大的复合物，这样的复合物与细胞表面接触不好，很难被吸收，那么机体又如何提高脂质水解产物的吸收效率呢？

如果你洗过油碗的话，就一定不会忘记为了把碗洗干净，会在水中加入少量去污剂。加去污剂的目的是为了将碗壁上的油珠分散到水相之中，实际上就是增加油的水溶性。脂质在消化道内的消化和吸收面临同样的问题，而解决问题的方法如出一辙，也是借助于机体自身合成的生物去污剂的“增溶”效应。

脂质在消化过程中使用的生物去污剂是由胆囊分泌的胆汁酸（盐）。胆汁酸（盐）作为胆固醇的衍生物，是一种两性分子，在肝细胞中合成以后，转移到胆囊，然后再分泌到消化道。当它与脂质接触以后，疏水区域与脂质结合，亲水区域则暴露在水相之中。于是，在胆汁酸（盐）的帮助下，大脂滴被分散成小脂滴，此过程称为乳化（emulsification）。乳化作用大大提高了水相和油相的界面面积，这对于增强水解酶的水解效率十分重要。

Quiz2 为什么胆囊被切除的“无胆”人士要少吃油性食物？

脂质被乳化以后，下一步就是脂肪酶、磷脂酶和胆固醇酯酶分别对脂肪、磷脂和胆固醇的水解。

脂肪的水解开始于胃，完成于小肠。参与脂肪水解的脂肪酶有 3 种：第一种是由舌组织分泌的舌脂肪酶（lingual lipase），它在胃的酸性 pH 下才有活性，在啮齿动物体内含量高，在人体内偏低；第二种是由胃分泌的胃脂肪酶（gastric lipase），它的最适 pH 是 4，切点在 sn-3 位置，对中链脂肪酸的水解快；第三种是胰脂肪酶（pancreatic lipase），它由胰腺大量分泌，与消化道内的绝大多数消化酶一样，以酶原的形式分泌，被胰蛋白酶激活。胰脂肪酶的活性还需要一种叫共脂肪酶（colipase）的辅助蛋白。共脂肪酶能够与脂肪酶形成稳定的复合物，并将脂肪酶锚定到水相与油相的界面。通过基因敲除技术得到的共脂肪酶被灭活的小鼠，刚刚出生的时候似乎很正常，但在两周以后，约有 60% 的小鼠死亡，活下来的小鼠对高脂肪食物中的脂肪吸收不良，体重很快下降。由此可见，共脂肪酶对脂肪的消化和吸收是必不可少的。脂肪在脂肪酶的作用下，产生甘油单酯和游离脂肪酸（FFA）。

Quiz3 溶血磷脂为什么具有溶血的效应？

目前市场上的许多减肥药，如奥利斯他（orlistat），实为胰脂肪酶的抑制剂，通过抑制脂肪在小肠内的消化和吸收而达到减肥功效。

消化道内水解磷脂的酶是由胰腺分泌的磷脂酶 A_2，它也是以酶原的形式分泌，被胰蛋白酶激活，需要 Ca^{2+} 的存在，水解产物为溶血磷脂和 FFA。

水解胆固醇酯的酯酶也由胰腺分泌，需要胆汁的激活，水解产物为胆固醇和 FFA。

二、脂质的吸收

脂质的吸收与单糖或氨基酸的吸收具有完全不同的机制，脂质消化的主要产物——甘油单酯和 FFA 通过简单扩散进入肠细胞，同时也有很多 FFA 通过质膜上特异性的转运蛋白进入肠细胞。

如图 21-3 所示，在脂肪酶作用下，单酰甘油（monoacylglycarol，MG）或甘油单酯和 FFA 作为产物被释放出来，但它们仍然与胆汁酸（盐）结合在一起，并与其他脂质的水解产物（如溶血磷脂和胆固醇）共同形成胶束（micelle）结构。胶束实际上就是脂质的各种水解产物与胆汁酸（盐）形成的微型聚合物（3~6 nm）。当胶束遇到肠细胞的刷状缘时，脂质（包括 MG、FFA、胆固醇和溶血磷脂）即被吸收。在这里，胶束充当了将脂质的消化产物从消化场所转移到吸收场所的“穿梭”工具。具体来说，胶束有两个功能：首先，它能及时移除脂肪消化的产物，使它们不能抑制脂肪酶的活性（产物抑制）；其次，它将不溶性的消化产物转运到细胞膜，使它们能够通过扩散的方式直接进入细胞。

脂质被吸收以后，脂肪和磷脂在肠细胞的光面内质网上重新合成，并与胆固醇、脂溶性维生素和载脂蛋白一起在高尔基体组装成乳糜微粒。随后，乳糜微粒通过胞吐的方式离开肠细胞，进入乳糜管，最后汇入血液。

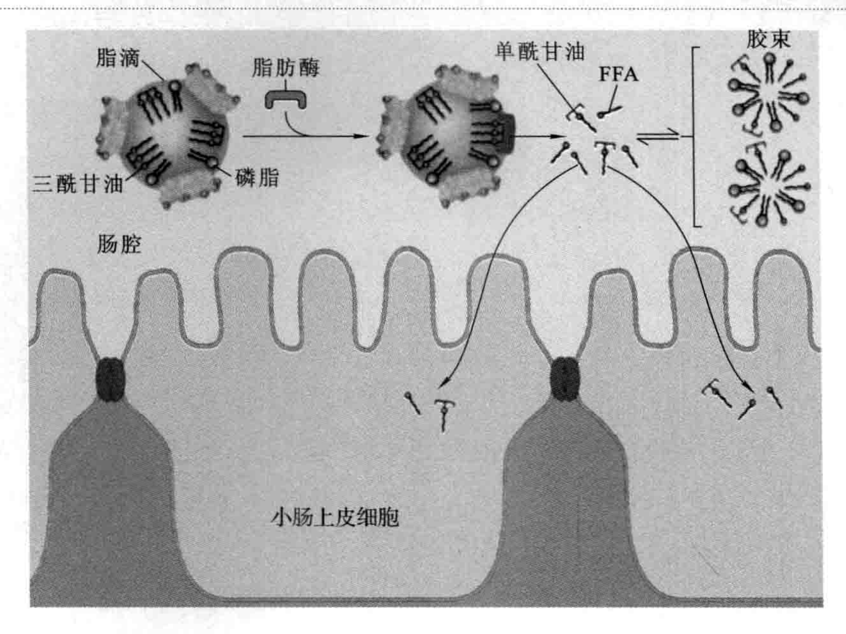


图 21-3 脂肪在消化道内的水解和吸收

第三节 蛋白质与核酸的消化和吸收

一、蛋白质的消化和吸收

膳食中的蛋白质，一般不能被肠道直接吸收，它们必须先被水解成氨基酸、二肽或三肽。肠道不能直接吸收完整蛋白质的原因有两个：首先，很少有蛋白质能够在大量的各种与膜结合的蛋白酶存在下不被水解；其次，正常的肠细胞质膜上没有专门负责跨膜转运蛋白质的转运蛋白，而且蛋白质也无法通过相邻肠道细胞之间的紧密连接。

然而有两个重要的例外。一是新生儿在出生后的一段时间内能够吸收完整的蛋白质，这一点非常重要，因为这可以使免疫系统还很脆弱的新生儿能从母乳中直接获得抗体。二是朊病毒能逃脱消化道内的蛋白酶水解并最终感染大脑。

(一) 蛋白质的酶促降解

消化道内参与蛋白质消化的蛋白酶由胃和胰腺分泌。其中，由胃分泌的是胃蛋白酶原，它受胃酸的激活而转变成有活性的胃蛋白酶；由胰腺分泌的有胰蛋白酶原、胰凝乳蛋白酶原、弹性蛋白酶原和羧肽酶原，它们受到胆囊收缩素的作用以后被释放；将胰蛋白酶原激活成胰蛋白酶的是由小肠黏膜分泌的肠激酶，将胰凝乳蛋白酶原、弹性蛋白酶原和羧肽酶原分别激活成胰凝乳蛋白酶、弹性蛋白酶和羧肽酶的是胰蛋白酶（参看第十二章“酶活性的调节”）。有趣的是，肠道内的蛋白酶也会因胰蛋白酶的作用而失活。

Quiz4 胃酸的产生与何种质子泵有关？胃酸有哪些功能？

蛋白质的消化开始于胃，由胃蛋白酶催化。胃蛋白酶的最适 pH 是 1.8，而如此低的 pH 只存在于胃里，因而当胃蛋白酶离开胃以后，活性就大大降低。

如图 21-4 所示，食物中的蛋白质经过胃和胰腺分泌的蛋白酶消化，在小肠内主要水解成寡肽。在小肠的刷状缘上，分布着各种肽酶。这些肽酶与前面提到的乳糖酶和麦芽糖酶一样，属于膜内在蛋白，而不是可溶性蛋白。它们能进一步水解肠腔内的寡肽，将寡肽转变成更小的小肽和游离的氨基酸。于是蛋白质消化的终产物（氨基酸、二肽和三肽）集中在肠细胞表面，为吸收做好了准备。

(二) 氨基酸、二肽和三肽的吸收

氨基酸被小肠细胞吸收的机制与单糖的吸收机制极为相似。在吸收细胞面向肠腔的质膜上，已

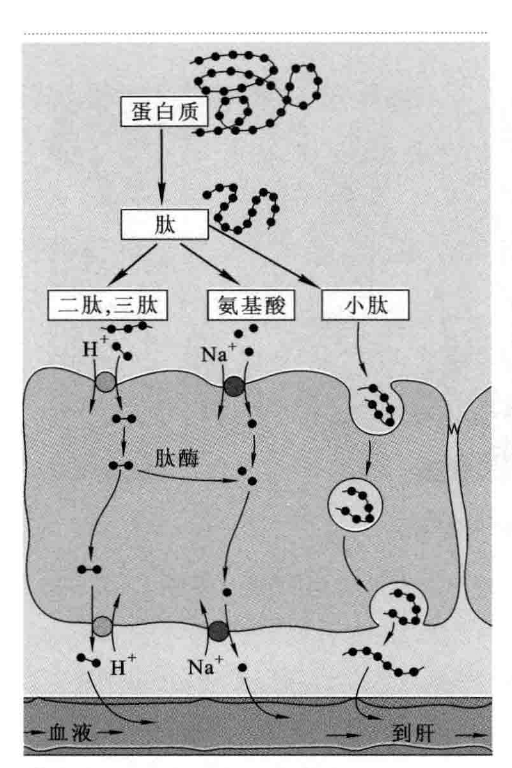


图 21-4　蛋白质的胞外水解和吸收

发现至少有4类依赖于 Na^+ 的氨基酸转运蛋白,分别转运酸性氨基酸、碱性氨基酸、中性氨基酸和脯氨酸(包括羟脯氨酸)。这些转运系统都属于次级主动运输。此外,还发现了不依赖于 Na^+ 的专门转运中性氨基酸、疏水氨基酸和碱性氨基酸的转运蛋白。

除了由特定的氨基酸转运蛋白来吸收氨基酸以外,在小肠上皮细胞还存在一种 γ-谷氨酰循环(γ-glutamyl cycle)。该循环也具有吸收氨基酸的功能,但需要谷胱甘肽。

e21-1 γ-谷氨酰循环

小肠上皮细胞几乎不能吸收长于3个氨基酸组成的寡肽,但却能有效地吸收二肽和三肽。与氨基酸吸收不同的是,二肽和三肽的吸收不需要 Na^+。一旦它们进入胞内,很快就被细胞质基质的肽酶水解成游离氨基酸。

被吸收的游离氨基酸和单糖分子一起,会在小肠上皮细胞产生一种驱动水吸收的渗透压梯度(osmotic gradient)。在小肠上皮细胞基底膜上,还含有将氨基酸转运出细胞而进入毛细血管的转运系统,但它们与 Na^+ 无关。

二、核酸的消化和吸收

从营养的角度来看,核酸的营养价值在几种生物大分子中是最低的,因此学习生化的人不该轻信所谓的核酸补品广告。食物中的核酸在消化道也会被消化和吸收。胰液中含有两类核酸酶,即核糖核酸酶和脱氧核糖核酸酶,它们被分泌到十二指肠分别水解RNA和DNA。但近来发现,植物来源的某些微RNA(micro RNA)能够在消化道被动物吸收,并被运送到特定的细胞,抑制胞内特定的基因表达。核苷酸带有负电荷,因此很难被小肠上皮细胞直接吸收,需要先被水解成核苷和磷酸,核苷还可再进一步水解成核糖(脱氧核糖)和碱基;这一吸收过程为主动转运。

第四节　真菌和食虫植物对生物大分子的消化和吸收

真菌虽不是像动物一类的消费者,但许多利用腐生性营养(saprophytic nutrition)。这意味着它们虽然不能摄入食物,但却能利用细胞外消化和吸收从环境中获取营养。

实际上,真菌能分泌多种消化酶,包括蛋白酶、脂肪酶和纤维素酶等,以便将周围的蛋白质、脂肪和多糖水解成为可被吸收的氨基酸、FFA、甘油和单糖等。

食虫植物(如猪笼草)也需要分泌一些水解酶(主要是蛋白酶),将捕获的动物进行消化和吸收。

框 21-1　身边的生物化学——猪笼草吃荤的秘密

食虫植物也称为食肉植物,是一种捕获并消化动物而获得营养的自养型植物。其大部分猎物为昆虫和节肢动物,偶尔可以捕食两栖动物、小型爬行动物或哺乳动物。被消化的动物可以为食虫植物提供必需的氮源,因为它们生长于土壤贫瘠,特别是缺少氮素的地区,例如酸性的沼泽和石漠化地区。

猪笼草就属于食虫植物中比较常见的一类。这一类食虫植物拥有一种独特的形如猪笼的吸取营养器官——捕虫囊(图 21-5)。猪笼草的捕虫囊内有蜜腺,能分泌蜜汁引诱昆虫,昆虫进入捕虫囊后,囊盖并不

闭合，但捕虫囊囊口内侧的囊壁很光滑，因此能防止昆虫爬出。捕虫囊下半部的内侧囊壁稍厚，并有很多消化腺，这些腺体分泌出稍带黏性的消化液储存在囊底；消化液与人体胃液很相似，具有消化昆虫的能力。



图 21-5 猪笼草的捕虫囊

至于消化液中究竟有哪些消化酶，这一直是一个谜。直到 2008 年，日本香川大学的 Tatsuro Hamada 使用蛋白质组学的研究方法，对猪笼草消化液中的蛋白质成分，进行了系统的研究。他们首先使用 SDS－聚丙烯酰胺凝胶电泳的方法，将消化液中的蛋白质进行分离，并通过银染的手段进行观测。随后，对观测到的蛋白质条带进行胶内的胰蛋白酶消化，再对得到的肽段通过液相色谱－质谱分析（liquid chromatography-MS），获得肽段的氨基酸序列，最后，将得到的序列与公共数据库上的蛋白质序列进行比对分析，最终发现了以下几种消化酶：①猪笼草蛋白酶（nepenthesin）Ⅰ和Ⅱ，该酶与胃蛋白酶相似，属于天冬氨酸蛋白酶；② β- 木糖苷酶（β-xylosidase）；③几丁质酶（chitinase）；④ β-1,3- 葡聚糖酶（β-1,3-glucanase）；⑤类奇甜蛋白（thaumatin-like protein）。在以上这些蛋白质中，只有猪笼草蛋白酶是用来消化猎物蛋白质的，几丁质酶可能既可以帮助猪笼草抗真菌，又能消化昆虫外骨骼的主要成分——几丁质，而 β- 木糖苷酶和葡聚糖酶的功能可能是参与捕虫囊的形成，类奇甜蛋白可能是由蜜腺分泌用来引诱昆虫上钩的。

科学故事——葡萄糖转运蛋白的发现

葡萄糖是生物体重要的能源，其进出细胞需要专门的转运蛋白，即葡萄糖转运蛋白(GLUT)。1952 年，Widder 及其同事系统研究了人胎盘细胞（placental cell）对葡萄糖的转运，发现葡萄糖转运具有饱和动力学的性质。于是，Widder 预测胎盘细胞转运葡萄糖需要一种载体蛋白。

基于红细胞没有内膜结构，且很容易得到，因此它的细胞膜一直被认为是研究物质跨细胞膜转运的首选材料。1977 年，Michihiro Kasahara 和 Peter C. Hinkle 终于从人的红细胞膜中，得到了第一种负责葡萄糖跨膜转运的蛋白质，从此开创了葡萄糖转运蛋白研究的新时代。Kasahara 和 Hinkle 使用的分离、鉴定方法如下：首先，使用低渗溶液处理红细胞，导致其细胞膜破裂，在去除血红蛋白以后得到血影。然后，利用去垢剂抽取膜上可能的 GLUT。在经过 DEAE- 纤维素柱层析以后，将分级得到的蛋白质与磷脂混合在一起，经超声波处理，得到重组有 GLUT 的脂质体。最后，使用稀释或透析的方法去除残留的去垢剂，并测定它们转运葡萄糖的活性。他们纯化出的葡萄糖转运蛋白是一种高度疏水、糖基化程度各异的蛋白质，后来被称为 GLUT1，以区分其他参与葡萄糖转运的蛋白质。GLUT1 的动力学特征与完整的红细胞基本相同。

1985 年，M. Mueckler 和 Harvey F. Lodish 等人利用 GLUT1 制备的兔抗血清，成功地从 HepG2 细胞的 cDNA 文库中克隆到人的 GLUT1 基因。根据 cDNA 的核苷酸序列，推导出 GLUT1 含有 492 个氨基酸残基。再通过疏水性分析，预测它含有 12 个跨膜的 α 螺旋。就在得到 GLUT1 的 cDNA 序列不久，科学家以此为探针，在其他组织中发现其他种类的 GLUT。到目前为止，至少已发现十多种与葡萄糖转运相关的蛋白质，它们分布在不同的组织，功能并不完全相同。

在以上各种 GLUT 分子中，GLUT4 尤为重要，因为它对胰岛素是敏感的。胰岛素能够迅速调节

GLUT4 在细胞膜的数目，从而影响细胞对葡萄糖的转运。

胰岛素调节 GLUT4 在细胞膜上数目的方式十分有意思。1999 年，K. A. Powell 等人使用基因工程的方法，构建了一种 GLUT4- 绿色荧光蛋白（GFP）的融合蛋白表达载体。GFP 是一种来源于海洋生物水母中的蛋白质，其多肽链中含有特殊的生色团结构，可在紫外线下发出稳定的绿色荧光。因此，科学家可利用 GFP 的这个性质，追踪它与其他蛋白质组成的融合蛋白在细胞中的定位。

Powell 将构建好的 GFP-GLUT4 融合表达质粒显微注射到 3T3-L1 脂肪细胞内，然后，观察胰岛素对 GLUT4 的分布有何影响。他们的结果表明：在没有胰岛素的时候，细胞内表达的 GFP-GLUT4 分布在胞内的囊泡膜上；在加入胰岛素以后，发现 GFP-GLUT4 重新分布，转而集中到细胞膜上。Powell 的实验证明，胰岛素可以通过调节 GLUT4 在细胞中的分布而改变葡萄糖的转运，从而来实现对葡萄糖代谢的影响。

目前，关于胰岛素调节葡萄糖转运的机制已比较清楚。有胰岛素时，胰岛素与其受体结合，促进细胞内含有 GLUT4 的囊泡与细胞膜融合，从而提高质膜上 GLUT4 的数目，致使葡萄糖转运速率加快；相反，在缺乏胰岛素的时候，位于质膜上的 GLUT4 因内吞作用而数目下降，最终使葡萄糖转运速率下降。

推荐网址：

e21-2 本章小结填填看

1. http://en.wikipedia.org/wiki/Carbohydrate_digestion（维基百科有关糖类消化的内容）
2. http://library.thinkquest.org/11226/main/c14txt.htm（一个在线的图书馆，提供有关蛋白质消化吸收的内容）
3. http://www.vivo.colostate.edu/hbooks/pathphys/digestion/smallgut/absorb_lipids.html（美国科罗拉多州立大学提供的有关脂质消化吸收的内容）

参考文献：

1. Zhang C Y, et al. Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA. *Cell Research*, 22: 107–126, 2012.
2. Hatano N, Hamada T. Proteome analysis of pitcher fluid of the carnivorous plant *Nepenthes alata*. *Journal of Proteome Research*, 7: 809–816, 2008.
3. Czech M P, Corvera S. Signaling mechanisms that regulate glucose transport. *J Biol Chem*, 274: 1 865–1 868, 1999.

第二十二章　糖酵解

糖酵解(glycolysis)是生物体最重要的分解代谢途径之一,它几乎发生在所有的活细胞中,只是速率有别。通过该途径,葡萄糖或者其他单糖在没有氧气的参与下被氧化成丙酮酸,并产生 NADH 和少量的 ATP;某些非糖物质(如甘油)也可以间接地进入此途径得到氧化分解。

本章将重点介绍糖酵解的发现、所涉及的十步反应、生理功能、调节机制以及终产物 NADH 和丙酮酸的代谢去向。

第一节　糖酵解的发现

对于糖酵解途径,真正意义上的研究可追溯到 1897 年,当时 Buchner 兄弟发现酵母发酵(fermentation)可以在破碎的细胞即酵母汁中发生。这个发现不仅否定了那个年代盛极一时的"活力"论(vital-force theory),更为重要的是,从此以后,人们可以直接在细胞的抽取物或无细胞系统(cell-free system)中研究代谢。

1905 年,A. Harden 和 W. Young 在酵母汁中发现无机磷酸盐是发酵反应所必需的成分。不久,Cyrus Hartwell Fiske 和 Yellagaprada Subbarow 在肌肉细胞中发现磷酸肌酸,Karl Lohmann 发现 ATP,Otto Meyerhof 提出热化学(thermochemistry)。这些发现共同促成了高能键概念的诞生和磷酸化反应的提出。很快,腺苷酸被认为是糖酵解的辅酶和能量的载体。

1929 年,Nilsson 偶然发现了 3- 磷酸甘油酸。与此同时,Kurt Lohmann 和 Fritz Lipmann 共同发现了 1,6- 二磷酸果糖能够转变成不易水解的磷酸酯(3- 磷酸甘油酸)。

1933 年,Gustav Embden 从肌肉细胞抽取物中分离得到 3- 磷酸甘油酸,将这种磷酸化合物重新放回到肌肉抽取物以后,发现它又立即转变成了丙酮酸和磷酸。

综合以上结果,Embden 提出了一个完整的糖酵解途径,只是其中的一些反应是在后来的研究中才得到证实的。鉴于 Embden 和 Meyerhof 在糖酵解研究中的杰出贡献,糖酵解又称为 Embden-Meyerhof 途径(Embden-Meyerhof pathway,EMP)。

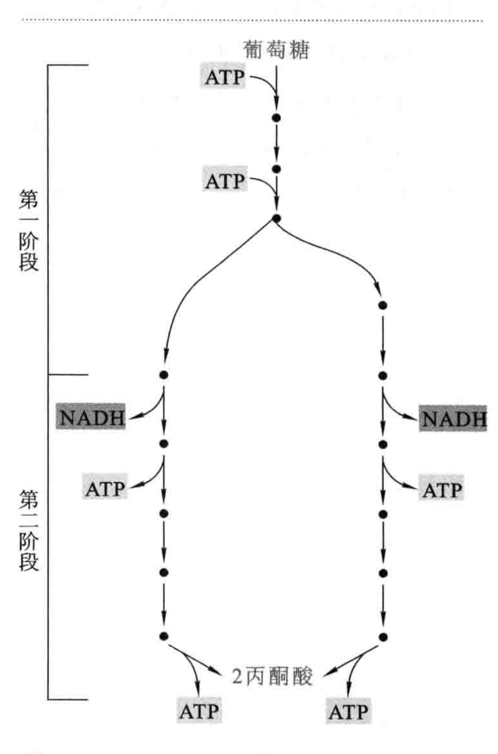


图 22-1　糖酵解的两阶段反应

第二节　糖酵解的全部反应

糖酵解发生在细胞质基质中,如果以葡萄糖为起始物质,整个糖酵解途径就由两个阶段十步反应组成(图 22-1,图 22-2):

(1) 第一阶段称为引发阶段(priming phase),有 5 步反应。在此阶段,1 分子葡萄糖经过 1,6- 二磷酸果糖转变成 2 分子 3- 磷酸甘油醛(glyceraldehyde-3-phosphate,

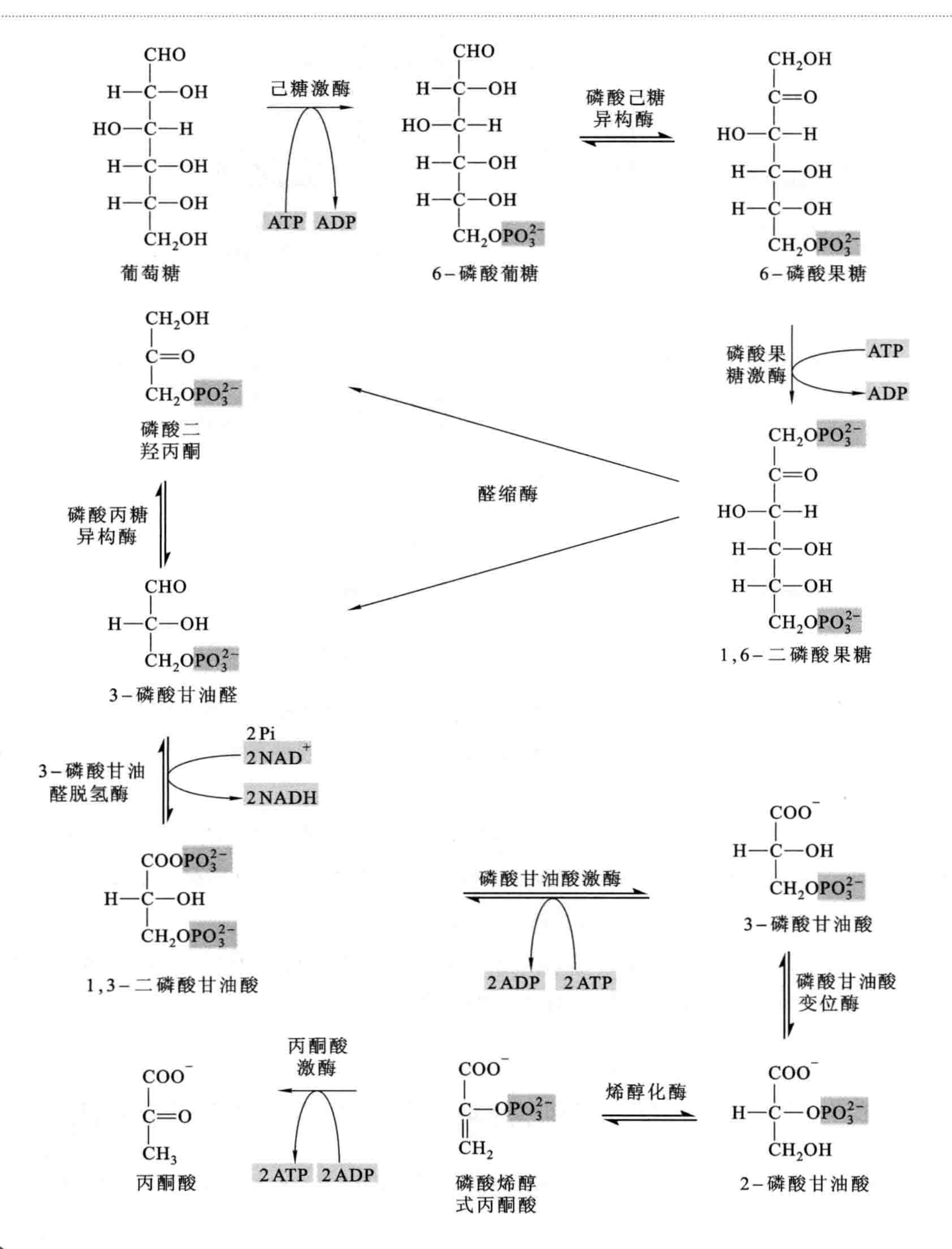


图 22-2 糖酵解的全部反应

GAP)，共消耗 2 分子 ATP。由于此阶段的反应消耗了 ATP，相当于能量货币的“投资”，因此有人称之为投资阶段（investing phase）。

(2) 第二阶段称为产能阶段（energy-yielding phase），也有 5 步反应。通过此阶段，2 分子 3- 磷酸甘油醛最终转变为 2 分子丙酮酸，同时产生 4 分子 ATP 和 2 分子 NADH。由于此阶段产生了 4 分子能量货币 ATP，在扣除第一阶段“投资”的 2 分子 ATP 以后，还“净赚”了 2 分子 ATP，因此第二阶段又称为收获阶段（harvesting phase）或获利阶段（profiting phase）。下面就分步介绍各反应。

1. 葡萄糖的磷酸化

这是糖酵解的第一步反应。在己糖激酶（hexokinase）或葡萄糖激酶（glucokinase）的催化下，葡萄糖 6 号位的羟基接受 ATP 分子上的 γ- 磷酸根，被磷酸化为 6- 磷酸葡糖（glucose-6-phosphate，G-6-P）。此反应的 $\Delta G^{\ominus\prime}$ 为较大的负值，这使其在细胞内正常的生理条件下为不可逆反应。与其他激酶一样，己糖激酶在催化反应中，巧妙地使用“诱导契合”机制，从而避免了 ATP 在活性中心发生

葡萄糖 ATP ADP Mg^{2+} 己糖激酶或葡糖激酶 6-磷酸葡糖

$\Delta G^{\ominus\prime} = -16.7$ kJ/mol

MgATP 己糖激酶或葡糖激酶 MgADP

葡萄糖 6-磷酸葡糖

水解的可能性(参看第八章“酶学概论”)。此外,反应还需要二价金属离子 Mg^{2+} 或 Mn^{2+}。这些二价金属离子的作用是与 ATP 形成复合物,屏蔽磷酸基团的负电荷,并且有助于酶的催化。缺乏 Mg^{2+} 的 ATP 将成为一种竞争性抑制剂。

葡萄糖的磷酸化至少有三方面的意义:一是降低了细胞内游离葡萄糖的浓度,这有利于胞外的葡萄糖通过 GLUT 进入胞内;二是葡萄糖由此带上负电荷,极性骤增,很难再从细胞中“逃逸”出去;三是葡萄糖由此变得不稳定,有利于其在细胞内的进一步代谢。

虽说己糖激酶和葡萄糖激酶都能催化这一步反应,但两种酶在底物的特异性、与葡萄糖的亲和力(K_m)、V_{max}、酶活性的调节以及基因表达上均有显著的差别(表 22-1)。

表 22-1　己糖激酶和葡萄糖激酶的比较

	己糖激酶	葡萄糖激酶
存在	几乎所有的细胞	肝细胞、胰腺的 β 细胞
底物特异性	葡萄糖、甘露糖、氨基葡萄糖、果糖、2- 脱氧葡萄糖等己糖	葡萄糖和 2- 脱氧葡萄糖
对葡萄糖的 K_m	0.1 mmol/L	10 mmol/L
V_{max}	低	高
产物反馈抑制	受 6- 磷酸葡萄糖反馈抑制	不受 6- 磷酸葡萄糖反馈抑制
基因表达	组成型酶	诱导酶

己糖激酶为组成型酶(constitutive enzyme),且对葡萄糖的亲和性较高,因此对于高等动物来说,其生理功能是确保细胞(即使在血糖浓度较低的情况下),尤其是脑细胞和红细胞能有持续的葡萄糖供应。己糖激酶对葡萄糖的高亲和性能够保证细胞内的葡萄糖分子及时被磷酸化,从而创造出有利于葡萄糖从血液进入细胞的浓度梯度。相比之下,葡萄糖激酶是一种诱导酶,与葡萄糖的亲和性较低,只有在血糖浓度处于较高的水平时,它才会被诱导表达并表现出活性,因此它的主要生理功能是及时清除饱餐以后血液里过多的葡萄糖,有助于稳定血糖水平。

2- 脱氧葡萄糖也可以作为这一步反应的底物,反应形成 2- 脱氧 -6- 磷酸葡萄糖,这种产物将是下一步反应的抑制剂。

2. 6- 磷酸葡糖的异构化

这是一步异构化反应,由磷酸己糖异构酶(phosphohexose isomerase)催化,也需要 Mg^{2+},反应的机

制涉及不稳定的烯二醇中间体。通过此反应,醛糖变成了酮糖,羰基从 1 号位变到 2 号位,这既为下一步磷酸化反应创造了条件,又有利于后面由醛缩酶催化的 C3 和 C4 之间的断裂反应。

6-磷酸葡糖 ⇌(Mg^{2+},磷酸己糖异构酶)6-磷酸果糖 $\Delta G^{\ominus\prime} = 1.7\ kJ/mol$

2- 脱氧 -6- 磷酸葡糖也能够与此酶的活性中心结合,但由于不能形成烯二醇中间体,因此无法完成反应,反而由于占据了活性中心而抑制酶的活性。

3. 磷酸果糖的激活

这又是一步消耗 ATP 的磷酸化反应,同样需要 Mg^{2+},$\Delta G^{\ominus\prime}$ 较大的负值使其为不可逆反应。催化这一步反应的酶通常为磷酸果糖激酶 -1(phosphofructokinase-1,PFK-1),有别于细胞内同样以 6- 磷酸果糖为底物的磷酸果糖激酶 -2(PFK-2),PFK-1 的产物为 1,6- 二磷酸果糖,而 PFK-2 的产物为 2,6- 二磷酸果糖。

这步反应的重要性在于,它是整个糖酵解途径的限速步骤(committed step),对糖酵解的调控主要是通过对此酶的调节而实现的(详见本章第六节"糖酵解的调节")。在植物、某些细菌、古菌和少数原生动物体内,还存在一种依赖焦磷酸的磷酸果糖激酶(pyrophosphate-dependent phosphofructokinase,PPi-PK),也能催化 6- 磷酸果糖转变成 1,6- 二磷酸果糖。但与 PFK-1 不同的是,此酶用焦磷酸代替 ATP,且催化的反应完全可逆。另外,某些古菌体内还有一种依赖于 ADP 的磷酸果糖激酶,也能催化 1,6- 二磷酸果糖的形成。

Quiz1 细胞有哪些方法可以产生 PPi?

6-磷酸果糖 + ATP →(Mg^{2+},磷酸果糖激酶-1)1,6-二磷酸果糖 + ADP $\Delta G^{\ominus\prime} = -14.2\ kJ/mol$

4. 1,6- 二磷酸果糖的裂解

这一步反应为可逆反应,根据逆反应,即醛醇缩合反应将催化此反应的酶命名为醛缩酶(aldolase)。通过该反应,1 分子六碳糖裂解成 1 分子磷酸二羟丙酮(dihydroxyacetone phosphate,DHAP)和 1 分子 3- 磷酸甘油醛(GAP)。

1,6-二磷酸果糖 ⇌(醛缩酶)磷酸二羟丙酮 + 3-磷酸甘油醛 $\Delta G^{\ominus\prime} = 23.8\ kJ/mol$

有两类醛缩酶(图 22-3):第一类来源于古菌和真核生物,为共价催化,在反应中,底物与活性中心的赖氨酸残基形成共价的希夫(Schiff)碱中间物;第二类主要来源于真菌和细菌,催化反应依赖活性中心结合的 Zn^{2+},为金属催化。

Quiz2 试设计一个实验,能够将这两种酶区分开来。

5. 磷酸丙糖的异构化

这是一步可逆的酮糖与醛糖互变的异构化反应。严格地说,它并不直接属于糖酵解途径,充其量不

第一类醛缩酶

醛缩酶肽链

亲核进攻

1,6-二磷酸果糖

共价希夫碱中间物

裂解

3-磷酸甘油醛

磷酸二羟丙酮

第二类醛缩酶

1,6-二磷酸果糖

3-磷酸甘油醛

图 22-3 醛缩酶的催化机理

过是将 DHAP 转变为糖酵解的中间代谢物 3-磷酸甘油醛，后者才可以进入下一步反应。但是，人们已习惯于将它视为第一阶段的最后一步反应。

催化该反应的酶为磷酸丙糖异构酶（triose phosphate isomerase，TIM），反应机制与第二步反应相似，其活性中心的 Glu 残基作为广义的碱行使催化。虽然从热力学的角度来看，此反应并不利于 3-磷酸甘油醛的形成，但由于 3-磷酸甘油醛很快被下一步反应所利用，因此，在细胞内此反应仍然能够顺利进行。

磷酸二羟丙酮 ⇌ 3-磷酸甘油醛 $\Delta G^{\ominus\prime} = 7.5\ \text{kJ/mol}$

TIM 所具有的一些动力学特征，以及在催化中反映出来的近乎完美的表现让很多科学家着迷。其极高的 k_{cat}/K_m 值（$2.4\times10^8\ \text{mol}^{-1}\cdot\text{L}\cdot\text{s}^{-1}$）表明它是一个接近完美的酶（near-perfect enzyme）（参看第九章"酶动力学"）。另外，它具有独特的防止副反应发生的机制（图 22-4），在反应中形成了磷酸烯二醇中间物，

磷酸二羟丙酮

烯二醇中间物

E：磷酸丙糖异构酶

3-磷酸甘油醛

图 22-4 磷酸丙糖异构酶的作用机理

这样的中间物如果离开酶分子，在溶液中很容易释放出磷酸根生成丙二醛，而异构化成3-磷酸甘油醛的机会并不多(前者反应速率比后者快近100倍)。但实际上，在细胞内形成丙二醛的可能性几乎为零，这是因为当烯二醇中间物形成以后，酶的构象发生变化，其分子上有一段由10个氨基酸残基组成的环，就像一个盖子，会堵住活性中心，致使烯二醇中间物无法离开酶分子，只能异构化成3-磷酸甘油醛。当3-磷酸甘油醛形成以后，封闭活性中心的环消失，产物得以释放。

6. 3-磷酸甘油醛的氧化及磷酸化

这是第二阶段的第一步反应，也是整个糖酵解途径中唯一的一步氧化还原反应，催化反应的酶为3-磷酸甘油醛脱氢酶(glyceraldehyde-3-phosphate dehydrogenase，GAPDH)，辅酶为NAD^+，产物为1,3-二磷酸甘油酸(1,3-bisphosphoglycerate，1,3-BPG)和$NADH/H^+$。在反应中，醛基变成了羧基，释放出来的能量一部分贮存在1,3-BPG的高能磷酸键上，还有一部分被NADH中的高能电子带走。

GAPDH由4个相同的亚基组成。每个亚基有个活性中心，而每个活性中心含有一个Cys残基和一个结合的NAD^+。Cys的巯基直接参与催化。以任意一个亚基为例(图22-5)：首先，酶活性中心的巯基S亲核进攻底物3-磷酸甘油醛分子上的羰基C，缩合成共价的半缩硫醛(thiohemiacetal)。随后，半缩硫醛被结合在酶分子上的NAD^+氧化成高能硫酯键，同时NAD^+被还原为NADH。此后，无机磷酸进攻高能硫酯键，并释放出另一种含有高能酰基磷酸键的混合型酸酐——1,3-二磷酸甘油酸。最后，溶液中的氧化型NAD^+置换出与活性中心结合的还原型NADH，促使酶恢复到原来的状态。

由此可见，整个酶促反应与活性中心Cys残基上的自由巯基密切相关，因此，任何破坏巯基的试剂，如碘代乙酸或甲基汞都是此酶的抑制剂(图22-6)，同时也是糖酵解的抑制剂。

虽然GAPDH的每一个亚基都可以独立地进行催化，但是4个亚基对NAD^+的结合呈现负协同效应，因此各个亚基对NAD^+的亲和力是不同的。GAPDH对NAD^+的结合具有负协同性，使其对[NAD^+]变化不敏感，这一性质让GAPDH在[NAD^+]较低的情况下也能照常进行催化(参看第九章“酶动力学”)。

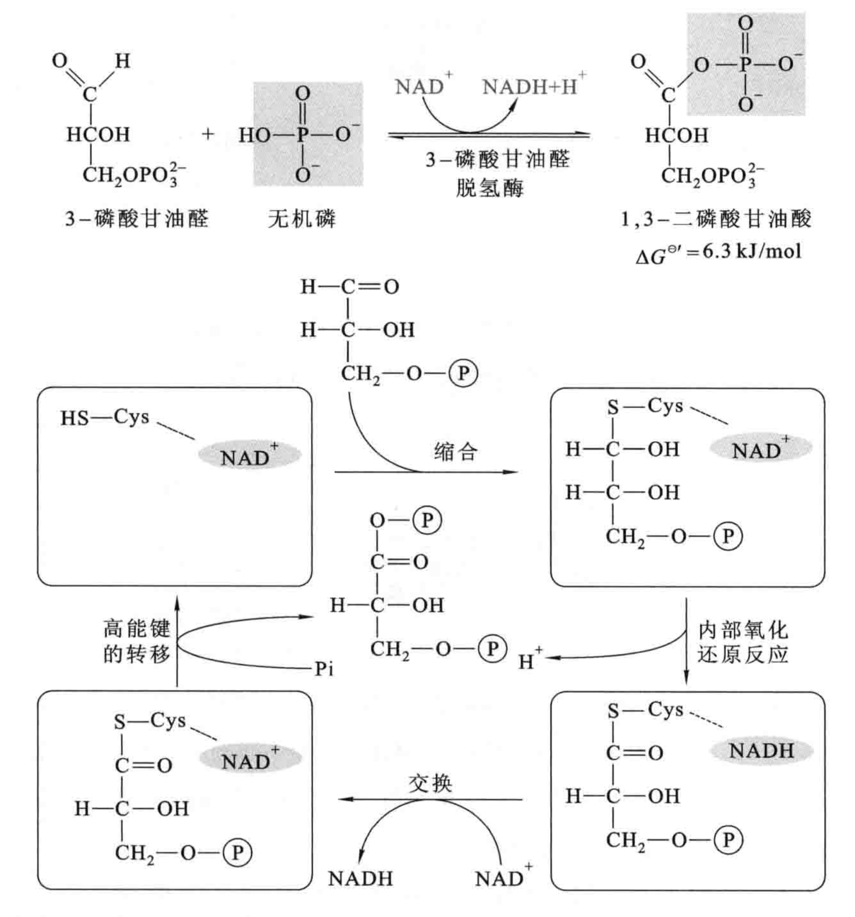


图22-5 3-磷酸甘油醛脱氢酶的作用机理

在化学结构和化学性质上，砷酸（arsenate，AsO_4^{3-}）与无机磷酸极为相似，因此可以代替无机磷酸参加反应，形成 1- 砷酸 -3- 磷酸甘油酸（1-arseno-3-phosphoglycerate），但这样的产物极不稳定，很容易自发地水解成 3- 磷酸甘油酸并产热，无法进入下一步底物水平的磷酸化反应，但是，砷酸并不抑制糖酵解，反而能刺激糖酵解以更快的速率进行。砷酸的这种作用与在第二十章“生物氧化”中提到的解偶联剂机制类似。1- 砷酸 -3- 磷酸甘油酸的自发水解，会导致 ATP 合成受阻，影响细胞的正常代谢，这也是砷酸有毒性的原因。

图 22-6 碘代乙酸和甲基汞抑制 3- 磷酸甘油醛脱氢酶活性的机理

7. 1，3- 二磷酸甘油酸的底物水平磷酸化

这是一步底物水平的磷酸化反应——先合成一种高能中间物，然后在一种转移酶的催化下，中间物上的高能键被转移给 ADP，形成 ATP。这里的高能中间物是 1，3- 二磷酸甘油酸，由磷酸甘油酸激酶催化，需要 Mg^{2+}。由于 1 分子葡萄糖能够产生 2 分子 1，3- 二磷酸甘油酸，因此这一步反应可合成 2 分子 ATP，正好与第一阶段投入的 2 分子 ATP 相抵消。尽管这一步反应的 $\Delta G^{\ominus\prime}$ 为较大的负值，但在细胞内的条件下，该反应仍然是可逆的。

$\Delta G^{\ominus\prime} = -18.5\ kJ/mol$

8. 3- 磷酸甘油酸的异构化

这又是一步可逆的异构化反应，由磷酸甘油酸变位酶（phosphoglycerate mutase）催化，也需要 Mg^{2+}。经过这一步反应，甘油酸上的磷酸基团从 3 号位变位到 2 号位。

$\Delta G^{\ominus\prime} = 4.4\ kJ/mol$

不同来源的磷酸甘油酸变位酶具有不同的催化机制，其中一类需要微量的 2，3- 二磷酸甘油酸（2，3-bisphosphoglycerate，2，3-BPG）作为辅助因子，并需要活性中心的一个 His 残基，比如来源于酵母细胞的变位酶（图 22-7）；另一类则不需要 2，3-BPG，其变位实际上是磷酸基团在 3- 磷酸甘油酸分子内部的转移，比如来自小麦麦胚细胞的变位酶。

Quiz3 根据酶的分类规则，来源于酵母的磷酸甘油变位酶应该属于六类酶中的哪一类？

2，3-BPG 由 1，3-BPG 变位而来，催化变位反应的酶是 1，3- 二磷酸甘油酸变位酶（1，3-bisphosphoglycerate mutase），该酶不仅催化变位反应，还具有磷酸酶的活性，因此它是一个双功能酶

3-磷酸甘油酸 + 酶-His-Ⓟ ⇌ 2,3-二磷酸甘油酸 + 酶-His

2,3-二磷酸甘油酸 + 酶-His ⇌ 2-磷酸甘油酸 + 酶-His-Ⓟ

总反应:3-磷酸甘油酸 ⇌ 2-磷酸甘油酸

图 22-7 依赖于 2,3- 二磷酸甘油酸的磷酸甘油酸变位酶的作用机理

(bifunctional enzyme)。磷酸酶活性将 2,3-BPG 水解成 3- 磷酸甘油酸。值得注意的是,如果 1,3-BPG 通过这种途径——甘油酸二磷酸支路(BPG shunt)转变成 3- 磷酸甘油酸,就跳过了一步底物水平磷酸化的反应,等于 1 分子葡萄糖少产生 2 分子 ATP。

2,3-BPG 作为辅助因子的作用是激活磷酸甘油酸变位酶,使其转变为具有活性的磷酸化形式。在活性中心上的 1 个 His 残基接受 2,3-BPG 的 1 个磷酸基团以后,磷酸甘油酸变位酶即被激活,随后按照图 22-7 所示的方式,催化 3- 磷酸甘油酸的变位反应。

Quiz4 存放较久的血液中的血红蛋白氧合曲线有无变化? 这样的血液能否使用?

体内绝大多数细胞内的 2,3-BPG 浓度很低,只有红细胞例外(在红细胞内约为 4 mmol/L,相当于血红蛋白的浓度),不过,2,3-BPG 在红细胞内的主要功能是调节血红蛋白与氧气结合的能力(参看第三章“蛋白质的功能及其与结构之间的关系”)。

红细胞内的 2,3-BPG 的浓度受到负反馈机制的调控。由于脱氧血红蛋白与 2,3-BPG 的亲和力高于氧合血红蛋白,因此,人体在缺氧的条件下(例如从低海拔到高海拔或心肺功能长期受损),脱氧血红蛋白的浓度增大,游离的 2,3-BPG 随之下降,此时 1,3- 二磷酸甘油酸变位酶受到的反馈抑制减弱,2,3-BPG 的合成相应增加,人体内 2,3-BPG 的水平随之上升。

9. 2- 磷酸甘油酸的烯醇化

这是一步由烯醇化酶(enolase)催化的反应,需要 Mg^{2+}。烯醇化酶的作用是促进甘油酸 -2- 磷酸上某些原子的重排,使能量富集到某一个化学键上,直接导致高能分子 PEP 的形成。氟化物能够与 Mg^{2+} 及磷酸基团形成络合物,干扰 2- 磷酸甘油酸与烯醇化酶的结合,使该酶的活性受到抑制。在测定血糖浓度之前,在血样中加入少量氟化钠以阻止糖酵解,这样可使测定的结果更准确。

2-磷酸甘油酸 ⇌（烯醇化酶，H_2O）磷酸烯醇式丙酮酸

$\Delta G^{\ominus\prime} = 7.5$ kJ/mol

10. 磷酸烯醇式丙酮酸的底物水平磷酸化

Quiz5 如果糖酵解的第六步反应用砷酸代替磷酸,那 1 分子葡萄糖经过糖酵解可以产生多少分子的 ATP?

这是糖酵解的最后一步反应,也是第二步底物水平的磷酸化反应和第三步不可逆反应,由丙酮酸激酶催化,需要 Mg^{2+} 和 K^+。通过此步反应,PEP 中的高能磷酸键转移给 ADP,糖酵解终于有了净的 ATP 合成。如果以 1 分子葡萄糖作为起始物质,到这一步反应结束,就可以净合成 2 分子 ATP。

磷酸烯醇式丙酮酸 + ADP $\xrightarrow[\text{丙酮酸激酶}]{Mg^{2+},K^{+}}$ 丙酮酸 + ATP

$\Delta G^{\ominus\prime} = -31.4\ kJ/mol$

丙酮酸（烯醇式） ⇌ 丙酮酸（酮式）

第三节　NADH 和丙酮酸的命运

糖酵解的终产物包括 ATP、丙酮酸和 NADH，产生的 ATP 可直接参与细胞内的各种需能反应，至于丙酮酸和 NADH 的去向，完全取决于细胞的类型和细胞内氧气的可得性。

一、在有氧状态下 NADH 和丙酮酸的命运

（一）NADH 的命运

在有氧状态下，NADH 在呼吸链被彻底氧化成 H_2O 并产生更多的 ATP。对于原核细胞，NADH 很容易进入位于质膜上的呼吸链，从而被氧化成 NAD^+，NAD^+ 得以再生并重新进入糖酵解。但对于真核细胞来说，糖酵解发生在细胞质基质，呼吸链则位于线粒体内膜，糖酵解产生的 NADH 是如何进入线粒体内膜上的呼吸链呢？ NADH 很容易直接通过通透性较好的外膜，但要通过线粒体内膜就必须借助于内膜上专门的穿梭系统（shuttle system）。

线粒体内膜上含有 3- 磷酸甘油穿梭系统（glycerol-3-phosphate shuttle system）和苹果酸 - 天冬氨酸穿梭系统（malate-asparate shuttle system）。它们在线粒体内膜上的相对比例因细胞种类而不同，例如脑细胞中主要是第一种，肝细胞中主要是第二种。

3- 磷酸甘油穿梭系统相对简单：首先，细胞质基质中的 3- 磷酸甘油脱氢酶将糖酵解产生的 NADH 氧化成 NAD^+，同时将磷酸二羟丙酮还原成 3- 磷酸甘油。随后，3- 磷酸甘油离开细胞质基质，进入线粒体，被内膜上的 3- 磷酸甘油脱氢酶重新氧化成磷酸二羟丙酮。磷酸二羟丙酮重新进入细胞质基质，脱下来的氢和电子交给 FAD 后，再通过 CoQ 进入呼吸链进一步氧化（图 22-8）。经过氧化磷酸化，原来的 1 分子 NADH 只能产生 1.5 分子的 ATP，损失了 1 分子 ATP。

苹果酸 - 天冬氨酸穿梭系统比较复杂（图 22-9）。与第一种穿梭系统不同的是，经苹果酸 - 天冬氨酸的穿梭，糖酵解产生的 NADH 变成了线粒体基质内的 NADH，再经过复合体Ⅰ进入呼吸链被彻底氧化。最后通过氧化磷酸化，1 分子 NADH 仍然可以产生 2.5 分子的 ATP，并无 ATP 的损失。

（二）丙酮酸的命运

在有氧状态下，丙酮酸通过线粒体内膜上的丙酮酸转运蛋白（pyruvate transporter），与质子一起进入线粒体基质，被基质内的丙酮酸脱氢酶系（pyruvate dehydrogenase complex）氧化成乙酰 -CoA。

1. 丙酮酸脱氢酶系的结构与组成

丙酮酸脱氢酶系是由丙酮酸脱氢酶（E_1）、二氢硫辛酸转乙酰酶（dihydrolipoyl transacetylase，E_2）和二氢硫辛酸脱氢酶（dihydrolipoyl dehydrogenase，E_3）通过非共价键结合在一起的稳定复合物。这三种酶在复合体中的数目和各自的辅助因子等参看表 22-2。

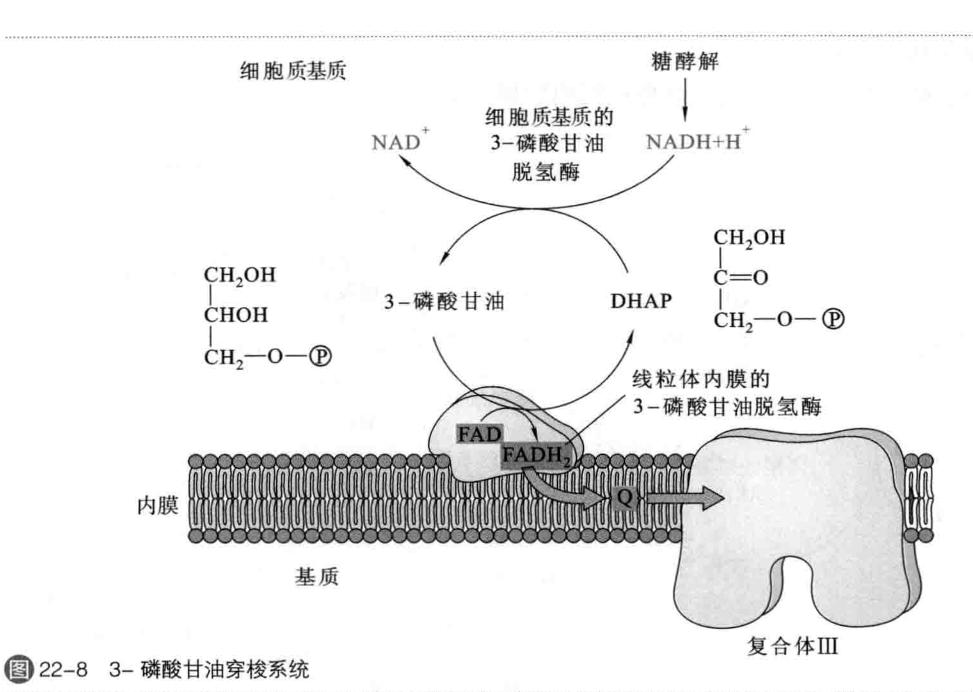


图 22-8 3-磷酸甘油穿梭系统

图 22-9 苹果酸 - 天冬氨酸穿梭系统

► 表 22-2 丙酮酸脱氢酶系的结构和组成

缩写	酶活性	亚基数目（个数）	辅助因子	维生素前体	辅助因子类型	催化的反应
E_1	丙酮酸脱氢酶	大肠杆菌 24、酵母 60、哺乳动物 20 或 30	TPP/Mg^{2+}	B_1	辅基	丙酮酸氧化脱羧
E_2	二氢硫辛酸转乙酰酶	大肠杆菌 24、酵母 60、哺乳动物 60	硫辛酰胺 CoA	硫辛酸 泛酸	辅基 辅酶	将乙酰基转移到 CoA
E_3	二氢硫辛酸脱氢酶	大肠杆菌 12、酵母 12、哺乳动物 6	FAD NAD^+	B_2 PP	辅基 辅酶	氧化型硫辛胺的再生

2. 丙酮酸的氧化脱羧

丙酮酸转变成乙酰 -CoA 的反应可分为 4 个阶段(图 22-10):

图 22-10 丙酮酸转变成乙酰 -CoA 的四步反应

(1) 丙酮酸的脱羧　由 E_1 催化,需要焦磷酸硫胺素(TPP)参与反应。TPP 的作用是作为碳负离子去亲核进攻丙酮酸的羰基碳,诱发脱羧基反应,形成羟乙基 TPP(图 22-11)。

(2) 碳单位的转移和氧化　由 E_2 催化,硫辛酰胺作为辅基。在 E_2 的催化下,氧化型硫辛酰胺上的二硫键被还原,TPP 上的羟乙基则被氧化成乙酰基,并被转移到硫辛酰胺的硫原子上形成高能硫酯键,同时 TPP 得以再生。

(3) 碳单位的再次转移　仍然由 E_2 催化,但需要 CoA 参与反应。在 E_2 的催化下,乙酰基从硫辛酰胺上的硫原子转移到 CoA 分子的巯基上,产生高能分子乙酰 -CoA,而原来的氧化型硫辛酰胺变成还原型硫辛酰胺。

图 22-11 丙酮酸脱氢酶的催化机理

(4) 氧化型硫辛酰胺的再生　由 E_3 催化,需要 FAD 和 NAD^+ 作为辅助因子。硫辛酰胺通过柔长的碳氢链系在 E_2 分子上(参看第十四章“维生素与辅酶”),因此它很容易在 E_2 和 E_3 活性中心之间进行转移。当还原型硫辛酰胺进入 E_3 的活性中心以后,随即被氧化,与 E_3 紧密结合的 FAD 先被还原成 $FADH_2$,再被 NAD^+ 氧化成 FAD,NAD^+ 则最终被还原成 NADH。

氧化型硫辛酰胺的再生对于丙酮酸脱氢酶系的持续运转十分重要,砒霜的主要成分亚砷酸能与还原型的硫辛酰胺形成共价的复合物,从而阻止氧化型硫辛酰胺的再生(图 22-12),这必然会导致整个酶系的失活。实际上,在三羧酸循环中也有一步类似的反应(由 α-酮戊二酸脱氢酶系催化,详见第二十三章“三羧酸循环”),因此,在细胞内,亚砷酸不仅能抑制丙酮酸转变成乙酰 -CoA,还能抑制三羧酸循环,可见其对细胞的毒性之大。

图 22-12 砒霜的毒性机理

二、在缺氧或无氧状态下 NADH 和丙酮酸的命运

细胞在缺氧或无氧的状态下，NADH 与丙酮酸的命运是相互关联的。虽然不同的生物所采取的手段不一定相同，但是目的完全一致，都是为了保证氧化型辅酶 I 即 NAD^+ 的再生。细胞质基质内辅酶 I 的量是一定的，如果没有特殊的机制让 NADH 重新转变成 NAD^+，那么，随着糖酵解的进行，一定时间以后，细胞质基质内的辅酶 I 都变成了还原型，糖酵解将会因 NAD^+ 的枯竭而自动停止，这必然导致灾难性的后果。

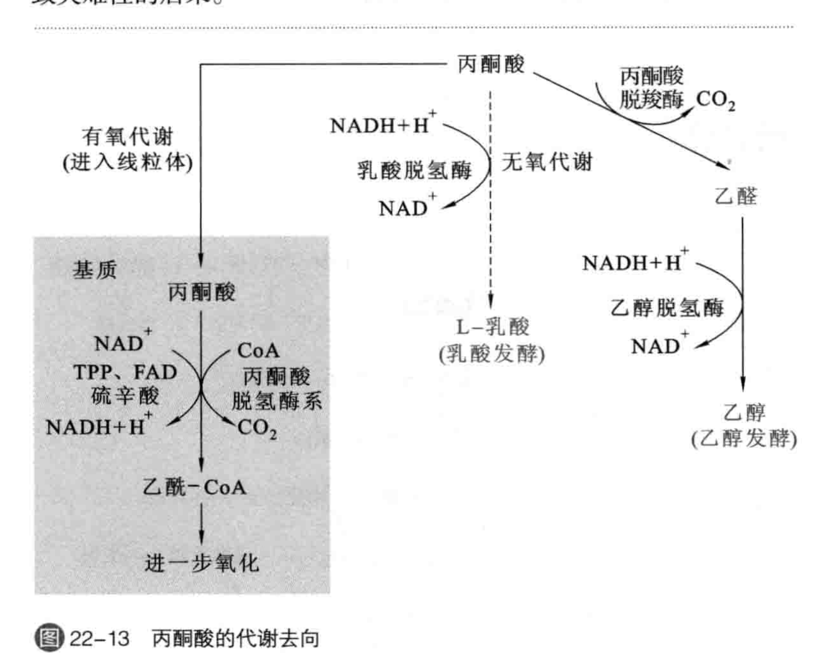


图 22-13 丙酮酸的代谢去向

在缺氧或无氧的条件下，最常见的再生 NAD^+ 的两种方法是乳酸发酵和乙醇发酵（图 22-13）。此外，在一些微生物体内，还有一些不常见的再生方式。这些方式不但解决了 NAD^+ 再生的问题，而且经常被人类应用于食品工业。例如，乳酸发酵用于制作酸奶和泡菜等，乙醇发酵用于酿酒、制作面包等，丙酸发酵用于制作奶酪。

（一）乳酸发酵

对于乳酸杆菌和其他一些微生物来说，再生 NAD^+ 的方法是使用丙酮酸作为氧化剂，在乳酸脱氢酶的催化下，将 NADH 重新氧化成 NAD^+，而丙酮酸则被还原成乳酸。

当高等动物的细胞出现氧债（oxygen debt）的情况下（例如肌肉剧烈运动），由于糖酵解产生的 NADH 来不及通过线粒体内膜上的穿梭系统重新转变成 NAD^+，这时体内也能发生乳酸发酵，但产生的乳酸需要被运输到肝细胞才可以重新被氧化成丙酮酸。而哺乳动物成熟的红细胞缺乏线粒体，因而每时每刻都在通过乳酸发酵再生 NAD^+。

（二）乙醇发酵

这种再生 NAD^+ 的方式也主要发生在许多微生物体内，如酿酒酵母等。具体过程：首先，在依赖于 TPP 的丙酮酸脱羧酶（pyruvate decarboxylase）催化下，丙酮酸脱羧变成乙醛。然后，乙醛在乙醇脱氢酶

框 22-1 生化趣事——不喝酒也醉酒的秘密

想象一下，吃了一碗面条或一包薯片后，你竟然醉了。真可谓酒不醉人人自醉也！

这样的事情你相信吗？但这正是二战结束后，驻扎在东京的美国大兵 Charles Swaart 经常发生的事情，虽然他一滴酒也没有沾过，但会完全进入醉酒的状态。在接下来的 20 年里，Swaart 不断地处在难以解释的醉酒状态。每一次醉酒都会给他的肝带来损害，年复一年，问题变得越来越严重。但巧合的是，在 1964 年，Swaart 听说有一个日本人，跟他有相同的遭遇。25 年后，一个内科医生终于诊断出了其中的病因。原来，在他们的肠道里，生活着一种突变的白色念珠菌，这些突变的酵母细胞利用摄入的糖类作为原料，来制造乙醇。这些酵母细胞中的代谢途径是糖酵解和酒精发酵。于是，在医生的建议下，Swaart 开始尝试服用几种抗生素来治疗。直到 1975 年，在他肠道内的所有突变酵母细胞都被杀死以后，他的生活才终于恢复正常。

为什么医生找到病因那么难呢？原因可能是，非发酵的白色念珠菌是人肠道里的常客，一般是想不到从肠道中去寻找能够产生乙醇的突变株的。那么这种突变株是怎么产生的呢？很有可能是广岛和长崎爆炸的原子弹引起核辐射造成的。

的催化下，被还原成乙醇，与此同时，NADH被氧化成NAD^+。人体虽有乙醇脱氢酶，但无丙酮酸脱羧酶，因此人体不能进行乙醇发酵，只能分解乙醇。

（三）其他形式的发酵

除了乳酸发酵和乙醇发酵以外，有些微生物还可以使用其他方式再生NAD^+。例如，一些细菌可进行丙酸发酵，即将丙酮酸还原成丙酸；再如，还有一些细菌进行丁二醇（butylene glycol）发酵而引起食物腐败。然而，无论是哪一种形式的发酵，实质上都是用NADH里面的电子去还原特定的有机分子。

第四节　其他物质进入糖酵解

前面是以葡萄糖作为起始物质来描述糖酵解的，实际上，细胞内还有一些物质也可以直接或间接进入糖酵解（图22–14），下面就选择其中的几种重要物质作简单介绍。

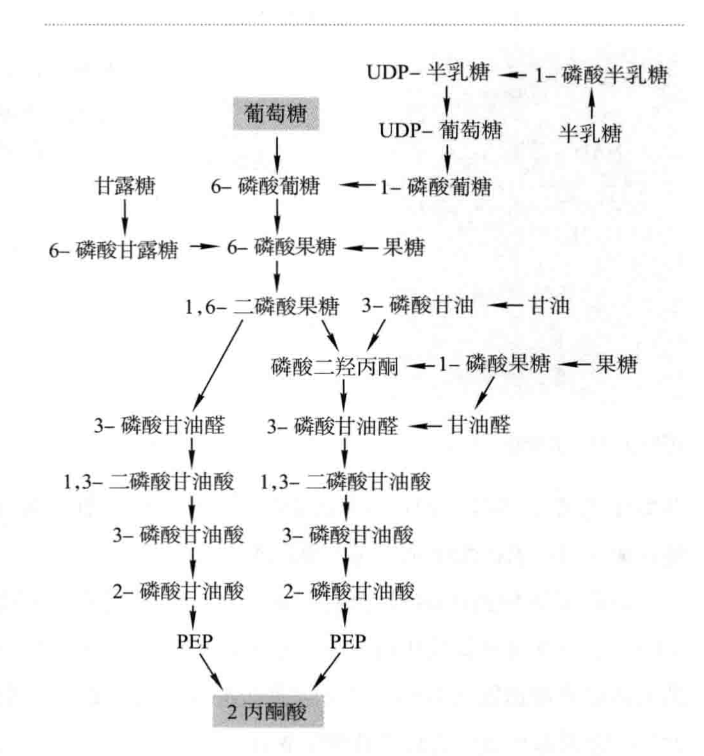


图22–14　甘油和其他单糖进入糖酵解的途径

（1）糖原　糖原的组成单位是葡萄糖，其中的葡萄糖单位主要经糖原磷酸化酶a的催化，被磷酸解成1–磷酸葡糖，1–磷酸葡糖经磷酸葡糖变位酶的催化，转变为6–磷酸葡糖而进入糖酵解。在催化机制上，磷酸葡糖变位酶与糖酵解中的磷酸甘油酸变位酶相似（图22–15），但磷酸化的氨基酸残基是Ser，而不是His。

Quiz6　如果进入糖酵解的葡萄糖是来自糖原的磷酸解，那么一分子葡萄糖可产生多少ATP？

（2）果糖　果糖可通过两种方式进入糖酵解，第一种是在己糖激酶的催化下（主要发生在肌细胞和肾细胞），形成6–磷酸果糖而进入；第二种是在果糖激酶的催化下（主要发生在肝细胞），转变为1–磷酸果糖，1–磷酸果糖在1–磷酸果糖醛缩酶（醛缩酶B）催化下裂解为磷酸二羟丙酮和D–甘油醛，前者经异构化变成3–磷酸甘油醛而进入糖酵解，后者在丙糖激酶（triose kinase）的催化下，形成3–磷酸甘油醛而进入糖酵解。

（3）甘露糖　甘露糖进入糖酵解的方法比较简单，先是在己糖激酶的催化下转变成6–磷酸甘露

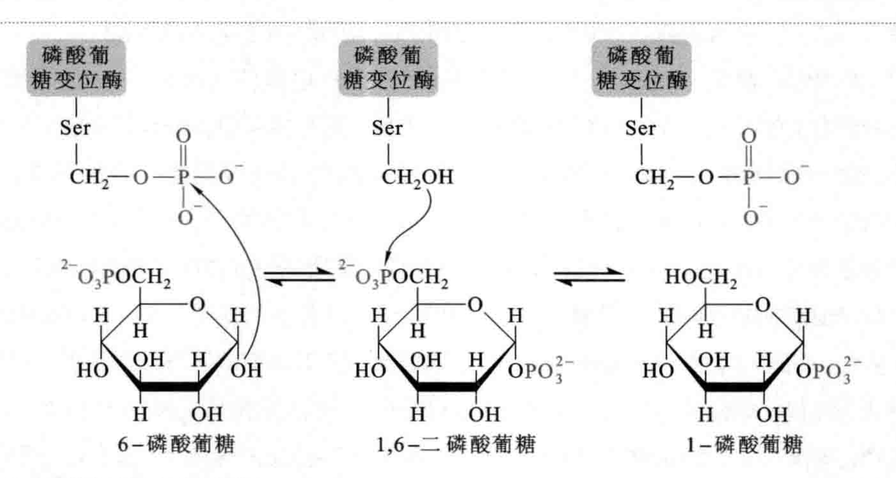


图22–15　磷酸葡糖变位酶的催化机制

糖，随后，在磷酸甘露糖异构酶的催化下，6- 磷酸甘露糖异构成 6- 磷酸果糖而进入糖酵解。

(4) 甘油　甘油虽不属于糖，但在细胞内也可以进入糖酵解。先是在甘油激酶(glycerol kinase)的催化下转变为 3- 磷酸甘油，然后在磷酸甘油脱氢酶(glycerol phosphate dehydrogenase)催化下，进一步转变为 DHAP，DHAP 经过磷酸丙糖异构酶的催化而进入糖酵解。

以上几种物质很容易进入糖酵解进行氧化分解，但与葡萄糖、果糖和甘露糖同属于己糖的半乳糖要进入糖酵解，却很复杂，需要通过一种很特殊的代谢途径——Leloir 途径(Leloir pathway)才能进入。

e22-1 Leloir 途径

第五节　糖酵解的生理功能

作为最古老的代谢途径，糖酵解几乎存在于所有的细胞，它对于维持细胞的正常功能是必不可少的，其主要的生理功能有以下几点：

(1) 产生能量货币 ATP　虽然糖酵解生成 ATP 的效率远远低于糖的有氧代谢，但是对于很多细胞来说却是主要的，甚至是唯一合成 ATP 的手段。例如，厌氧生物、无氧状态下的兼性生物(facultative organism)和哺乳动物成熟的红细胞，它们都以糖酵解作为产生 ATP 的唯一途径。而体内的某些组织，在缺氧的情况下(如肌肉组织)，或者因线粒体的数目有限(如视网膜和睾丸组织等)，会以糖酵解作为合成 ATP 的主要途径。

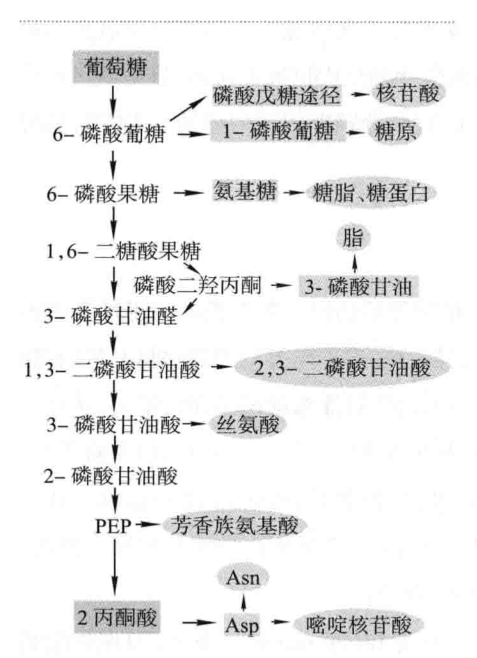


图 22-16　糖酵解某些中间物的代谢流向

(2) 为细胞内其他物质的合成提供原料　糖酵解的许多中间物可以离开糖酵解，作为细胞合成其他物质的前体(图 22-16)。例如，丙酮酸可直接为合成 Ala 提供碳骨架(参看第三十一章“氨基酸代谢”)，DHAP 可作为合成甘油的原料(参看第二十八章“脂肪、磷脂和糖脂的代谢”)，而 1,3-BPG 是 2,3-BPG 的直接前体，6- 磷酸葡糖则是糖原合成的前体(详见第二十七章“糖原代谢”)。

(3) 肿瘤与糖酵解　体内肿瘤细胞的生长和分裂速率极快，对氧气的需要通常会超过血管的供氧能力。随着肿瘤的生长，肿瘤细胞获取氧气的能力减弱，便会处于缺氧(hypoxia)的状态。在这种状态下，糖酵解成为细胞获得 ATP 的主要来源。此时，一种受缺氧诱导的转录因子(hypoxia inducible transcription factor，HIF-1)被激活，并进入细胞核，与 DNA 上的缺氧应答元件(hypoxia response element)结合，从而诱导绝大多数参与糖酵解的酶和参与葡萄糖跨膜转运的 GLUT1、GLUT2 的表达，这就通过“量变”的方式大大提高了糖酵解的速率，从而满足肿瘤细胞对 ATP 的需要。肿瘤细胞这种对缺氧环境的适应能力，使之能够生存下来一直到新血管的增生。事实上，HIF-1 还可以同时诱导促进新血管生长的血管内皮生长因子(vasular endothelial growth factor，VEGF)的表达。

(4) 糖酵解途径中少数酶的兼职功能　在第三章有关蛋白质功能的描述中，已介绍了兼职蛋白的概念并列举了许多重要的例子，其中就提到了几种与糖酵解有关的酶，如磷酸己糖异构酶可兼作神经白介素和自分泌运动因子，3- 磷酸甘油醛脱氢酶可在细胞核中兼作尿嘧啶 -DNA 糖苷酶，参与 DNA 的碱基切除修复(参看第三章“蛋白质的功能及其与结构之间的关系”)。事实上，还有一些与糖酵解酶有关的兼职蛋白，如己糖激酶Ⅱ和烯醇化酶等。

e22-2 糖酵解途径中的其他兼职蛋白

糖酵解作为一条维持细胞基本活动必需的代谢途径，在完成它的“本职工作”之外，竟然有近一半

的酶能够去“兼职”，即离开糖酵解途径，去行使其他功能。这说明随着生物的进化，一些基因的编码能力得到加强，“旧基因”和“老蛋白”完全能够挖掘出新功能，生物的每一项新功能不一定要通过产生新的基因来实现。这同时也提醒每一位从事生命科学研究的人，那些看起来似乎已经研究透的内容照样也能带来新的发现和惊喜！

第六节　糖酵解的调节

细胞内的任何代谢途径都是受到严格调控的，糖酵解也不例外。除了控制底物，即葡萄糖的可得性以外，机体主要通过调节几个关键限速酶的活性来控制糖酵解。那么糖酵解有哪几步反应是由限速酶催化的呢？糖酵解共有 10 步反应，只有 3 步反应为不可逆反应(分别为第一步、第三步和最后一步)，其余 7 步反应都是可逆反应。可逆反应为糖酵解和糖异生共有，只不过反应方向相反。显然，细胞不会选择 7 步可逆反应作为调控位点，若这 7 步可逆反应受到抑制或者激活，那糖酵解和糖异生也就同时受到抑制或激活，这样的结果显然是不合适的。既然糖酵解和糖异生分别属于糖的分解代谢和合成代谢，那么最好的调控策略应该是在糖酵解活跃的时候，糖异生恰好被抑制，反之亦然。由此看来，糖酵解的限速酶就是催化那三步不可逆反应的酶，即己糖激酶(包括葡萄糖激酶)、PFK-1 和丙酮酸激酶，其中 PFK-1 是糖酵解最重要的调节位点。那么，为什么机体偏偏选择第三步反应作为糖酵解最重要的调节位点呢？这是因为，第一步反应同时也是糖代谢另外一条途径——磷酸戊糖途径的第一步反应(详见第二十四章“磷酸戊糖途径”)，而由丙酮酸激酶催化的反应位置太靠后，只有第三步反应作为最重要的调节位点最合适。下面就主要以高等动物体为例，分别介绍葡萄糖的可得性对糖酵解的调节以及各限速酶活性的调节机制。

一、葡萄糖的可得性

葡萄糖是糖酵解的起始底物，其水平高低将直接影响到糖酵解的活性。葡萄糖的可得性受细胞膜上的 GLUT 调节。迄今为止，已发现十多种不同的 GLUT，它们存在于不同类型的细胞膜上(参看第二十一章“生物大分子的消化和吸收”中的科学故事)。不同的 GLUT 对葡萄糖的亲和力不同，这里可以借用 K_m 来表示 GLUT 对葡萄糖的亲和力。显然，葡萄糖进入细胞的速率一方面取决于 GLUT 的类型，另一方面还取决于 GLUT 的数目。血糖的平均浓度约 5 mmol/L，但饱餐后，可达到 12 mmol/L。几乎所有的细胞都含有一种共同的 GLUT，其对葡萄糖的 K_m 约为 1 mmol/L，低于血糖的平均浓度。因此，大多数细胞摄入葡萄糖的速率相对恒定，并不受血糖浓度波动的影响。

然而，肝细胞和胰腺 β 细胞含有一种特别的 GLUT，其 K_m 高达 15~20 mmol/L。因此，这两种细胞摄入葡萄糖的速率与血糖浓度成正比。这对于肝细胞和胰腺 β 细胞的生理功能是极为重要的，一方面使胰腺 β 细胞能直接监测出血糖水平，以此来调节胰岛素的分泌，另一方面，也保证了在人体血糖高的情况下，肝细胞能快速摄入葡萄糖，以将其转变成糖原贮备起来，在饥饿的时候，再将葡萄糖转运给其他组织，特别是脑细胞。

肌肉细胞和脂肪细胞表达第三类 GLUT，主要是 GLUT4，其对葡萄糖的 K_m 约为 5 mmol/L。与其他 GLUT 不同的是，GLUT4 在膜上的数目受到胰岛素的快速调节(参看第十七章“激素及其受体介导的信号转导”)。胰岛素的分泌能促进它们从胞内的小泡膜移位到细胞膜上，使肌肉细胞和脂肪细胞在血糖偏高的情况下，能吸收更多的葡萄糖，并分别将其转化成肌糖原和脂肪贮存起来。

二、己糖激酶和葡萄糖激酶的调节

调节己糖激酶活性的主要方式是反应产物即 6- 磷酸己糖对它的反馈抑制。当糖酵解在其他位

点（主要是第三步反应）受到抑制的时候，6- 磷酸己糖的利用受阻，导致它在细胞内堆积，多余的 6- 磷酸己糖作为反馈抑制剂能进一步减弱糖酵解作用。

与己糖激酶不同的是，葡萄糖激酶并不受到 6- 磷酸己糖的抑制，但作为一种诱导酶，葡萄糖激酶在肝细胞内的浓度受到胰岛素的控制。饱餐以后，体内的血糖浓度骤升，这时胰岛素的分泌量增多。胰岛素一方面能诱导肝细胞膜上 GLUT 的表达，使血液中大量的葡萄糖分子进入肝细胞（降低血糖浓度）；另一方面还能诱导葡萄糖激酶基因的表达，使大量涌入肝细胞的葡萄糖能够迅速被磷酸化为 6- 磷酸葡糖。形成的 6- 磷酸葡糖绝大多数并不进入糖酵解，而是用来合成糖原，以便将细胞内多余的燃料及时贮存起来。由此可见，葡萄糖激酶不受 6- 磷酸己糖的反馈抑制，其作用在于帮助机体在葡萄糖充裕的情况下，及时将它们贮存起来，以备不时之需。

三、PFK-1 的调节

PFK-1 作为糖酵解最重要的限速酶，“把守”着糖酵解途径的入口，直接控制进入糖酵解途径的葡萄糖和其他几种己糖的流量，因此机体对它的调控是最复杂的，同时也是最精妙的。

PFK-1 是一个复杂的多亚基蛋白，调节其活性的主要手段是别构调节。有两组别构效应物能够改变它的活性，一组为负别构效应物，包括 ATP、柠檬酸和质子（低 pH），它们能够抑制 PFK-1 的活性；另一组为正别构效应物，包括 AMP、ADP 和 2,6- 二磷酸果糖（F-2,6-BP），它们能够激活 PFK-1 的活性。细胞选用这些物质作为 PFK-1 的别构效应物的意义在于：①对细胞的能量状态或能荷 $\left(\frac{[\mathrm{ATP}]+1/2\,[\mathrm{ADP}]}{[\mathrm{ATP}]+[\mathrm{ADP}]+[\mathrm{AMP}]}\right)$ 迅速作出反应；②对细胞内环境酸度的变化作出反应（质子）；③对细胞内替代性燃料（如脂肪酸和酮体）的使用作出反应（柠檬酸）；④对血液中胰岛素和胰高血糖素的比例变化做出反应（F-2,6-BP）。下面就对这几种别构效应物的具体机制加以讨论。

（一）ATP 的别构抑制

ATP 对 PFK-1 的别构抑制反映了在细胞处于较高的能量状态（[ATP]较高）时，理应将产能途径“降温”，从而使细胞的合成代谢活跃起来。但问题是 ATP 如何既能作为 PFK-1 的底物，又作为它的抑制剂呢？原来，PFK-1 分子上有两个 ATP 结合位点，一个是酶的活性中心，另一个是别构中心。它们与 ATP 的亲和力不一样，前者的亲和力比较高，后者则比较低。当细胞的能荷较低时，ATP 只会与 PFK-1 的活性中心结合，别构中心并无 ATP 结合，这时绝大多数酶处于 R 态，PFK-1 与底物 F-6-P 的动力学曲线为双曲线（图 22-17）。糖酵解以正常的速率进行，以使细胞产生足够的 ATP；相反，当细胞的能荷较高时，ATP 就有机会与酶的别构中心结合。一旦两者结合，酶的构象就发生变化，迅速从 R 态变为 T 态，这时酶与 F-6-P 的结合曲线大幅度地向右倾斜，变成 S 形，于是酶对 F-6-P 的亲和力降低，其活性自然就受到了抑制。

Quiz7 如果测定 PFK-1 对 ATP 的动力学曲线，预测曲线的形状是什么？如果在反应系统中加入大量的 AMP，这会对动力学曲线有何影响？

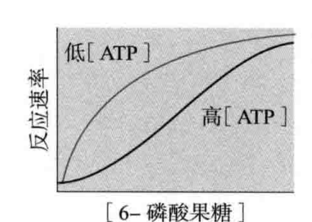


图22-17 ATP 浓度对 PFK-1 的活性影响

氧气对糖酵解的抑制效应——巴斯德效应（Pasteur effect）与 ATP 的抑制作用有关：在有氧的时候，生物体利用有氧代谢（三羧酸循环和氧化磷酸化）产生大量的 ATP，高水平的 ATP 作为负别构效应物，通过抑制 PFK-1 而抑制糖酵解。

（二）柠檬酸的别构抑制

柠檬酸作为三羧酸循环的中间物，实际上也是一种细胞能量状态的“指示剂”。当细胞内其他燃料（例如脂肪酸或酮体）大量氧化，并经过三羧酸循环和氧化磷酸化（参看第二十九章“脂肪酸代谢”）合成大量 ATP 而导致细胞能荷大幅度提高时，三羧酸循环将被削弱（详见第二十三章“三羧酸循环”），柠檬酸因此而积累，并通过线粒体内膜上的转运蛋白离开线粒体进入细胞质基质，作为负别构效应物抑制 PFK-1 的活性。柠檬酸似乎在向 PFK-1 传递一种信号：“细胞已经使用其他替代性燃料合成了大量的 ATP，你作为糖酵解途径中最重要的调节酶也该歇歇了”。事实上，柠檬酸在细胞质基质中还能激活脂肪酸合成的限速酶——乙酰 -CoA 羧化酶，这可以让脂肪细胞在富能的时候把多余的能量转化成脂肪贮存起来。

（三）质子的抑制

丙酮酸是糖酵解的终产物，可以经过乳酸发酵合成乳酸。哺乳动物的红细胞和肌细胞都可以进行乳酸发酵，但产生的乳酸需要经过血液循环到达肝细胞，并通过糖异生转变成葡萄糖，或者在心肌细胞内重新氧化成丙酮酸。在乳酸离开红细胞或肌细胞的时候，有质子伴随它被释放到血液里。因此，如果乳酸发酵过于强烈，将会导致血液酸度的提高，甚至出现酸中毒。使用质子作为反馈抑制剂就可以防止这种情况的出现。

（四）AMP 和 ADP 的别构激活

AMP 和 ADP 也是细胞能量状态的“指示剂”。当细胞内的能量状态较高时，它们的浓度就很低，这时很难作为 PFK-1 的正别构效应物起作用；相反，当细胞内的能量状态较低时，[AMP]和[ADP]就比较高，这时能够作为 PFK-1 的正别构效应物激活该酶的活性，使更多的底物进入糖酵解途径氧化分解，以改善细胞的能量状态。与 ADP 相比，AMP 的效果更佳，细胞内其他一些与能量代谢有关的酶也多用它作为别构效应物，这主要是因为在细胞内有反应：$2ADP \xrightleftharpoons{\text{腺苷酸激酶}} ATP+AMP$（参看第十九章“生物能学”），通过此反应，细胞内的 ADP 很容易转化为 ATP 和 AMP，这使得 AMP 的浓度进一步提高，相当于放大了 AMP 的量。显然，细胞选用 AMP 作为能量状态的“指示剂”，更有利于机体对细胞的能量状态做出既迅速又敏感的反应。

（五）F-2,6-BP 的别构激活

F-2,6-BP 被视为 PFK-1 最重要的正别构效应物，有实验表明，当哺乳动物细胞内缺乏它时，糖酵解将难以进行。合成 F-2,6-BP 的反应由磷酸果糖激酶 -2（PFK-2）催化，需要 Mg^{2+}。此反应为不可逆反应。

F-6-P 被磷酸化成 F-2,6-BP 以后，就意味着它进入了代谢终端，因为后者在细胞内不能被进一步地代谢。但在 2,6- 二磷酸果糖磷酸酶（fructose-2,6-bisphosphatase，F-2,6-BPase）催化下，F-2,6-BP 可以重新转变为 F-6-P。值得称奇的是，PFK-2 和 F-2,6-BPase 居然是同一个蛋白质的不同形式，前者为去磷酸化形式，后者为磷酸化形式（图 22-18）。

图 22-18　PFK-2 和 F-2,6-BPase 的活性调节和相互转变

Quiz8　若是将 PFK-2 的基因敲除，糖代谢将会受到什么影响？

PFK-2 或 F-2,6-BPase 由两个相同的亚基组成，每一个亚基含有 3 个结构域，即激酶结构域、磷酸酶结构域和调节结构域。调节结构域位于 N 端，磷酸酶结构域位于 C 端，激酶结构域位于中央。调节结

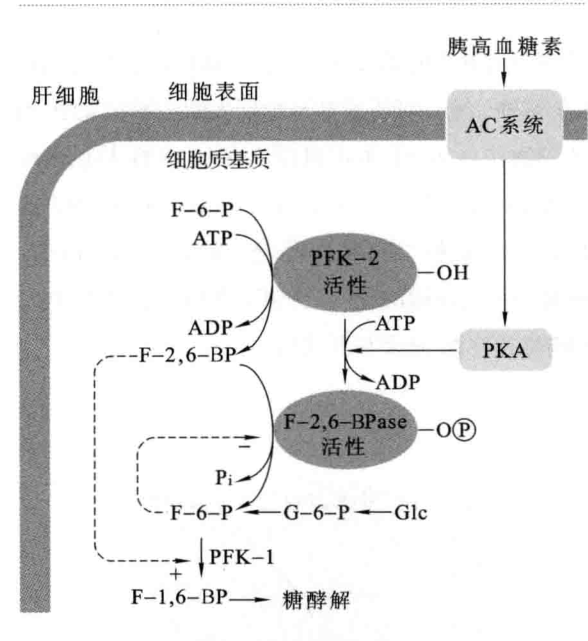


图 22-19 胰高血糖素对糖酵解的影响

构域是酶的磷酸化位点，它所处的状态（磷酸化或去磷酸化）直接决定了酶采取哪一种活性形式。控制调节结构域磷酸化和去磷酸化的两种激素则分别是胰高血糖素和胰岛素。

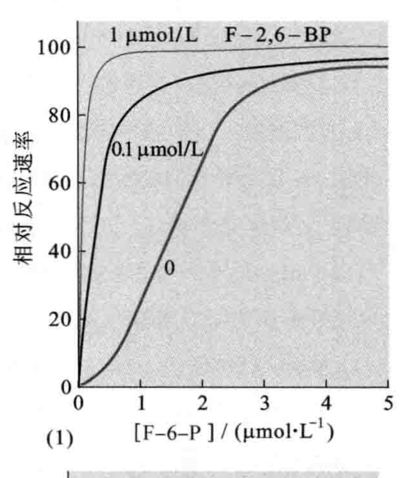
如图 22-19 所示，当血糖浓度下降时，胰高血糖素开始分泌，胰岛素停止分泌，胰高血糖素随血液循环到达肝细胞，并与位于肝细胞膜表面的受体结合，通过 AC 系统激活肝细胞内的 PKA（参看第十七章“激素及其受体介导的信号传导”），PKA 催化 PFK-2 在调节结构域的磷酸化，从而关闭其激酶活性，同时激活其磷酸酶活性。F-2,6-BPase 活性被激活以后，将 F-2,6-BP 水解成 F-6-P，PFK-1 的活性因缺乏 F-2,6-BP 的激活而受到抑制，糖酵解也因此受到抑制，与此同时，细胞内糖异生和脂肪酸的氧化途径则被激活（参看第二十五章“糖异生”和第二十九章“脂肪酸代谢”）。在被胰高血糖素作用后，肝细胞发生的种种代谢变化的目的只有一个，那就是阻止葡萄糖的分解，促进葡萄糖的合成，并最终提高血糖的浓度。

相反，如果血糖浓度上升，则胰高血糖素分泌停止，胰岛素开始分泌，胰岛素随血液循环到达肝细胞，通过肝细胞膜表面的 RTK 系统（参看第十七章“激素及其受体介导的信号传导”），激活肝细胞内特殊的磷蛋白磷酸酶，被激活的磷蛋白磷酸酶催化 F-2,6-BPase 在调节结构域的去磷酸化，从而关闭其磷酸酶的活性，同时激活其激酶的活性。PFK-2 活性被激活以后，将 F-6-P 磷酸化成 F-2,6-BP，PFK-1 的活性因 F-2,6-BP 的结合而受到激活，糖酵解也因此受到激活。与此同时，细胞内糖异生和脂肪酸的氧化则受到抑制，糖原和脂肪酸的合成加强。在被胰岛素作用后，肝细胞发生的种种代谢变化的目的也只有一个，那就是促进葡萄糖的分解和糖原合成，阻止葡萄糖的合成，并最终降低血糖的浓度。

由此看来，F-2,6-BP 是合成还是水解完全由血液中胰岛素和胰高血糖素的比例决定，而 F-2,6-BP 在合成和水解之间的转换使得机体能够对血糖浓度的变化迅速作出反应。

F-2,6-BP 对 PFK-1 活性的调节参看图 22-20(1)，在没有 F-2,6-BP 的时候，PFK-1 对 F-6-P 的动力学曲线为 S 形，这时酶主要以 T 态存在，与 F-6-P 的亲和力较低；在加入 F-2,6-BP 以后，PFK-1 对 F-6-P 的动力学曲线向左移动，变成了双曲线，并且[F-2,6-BP]越高，曲线移动的幅度就越大，这说明 F-2,6-BP 作为别构效应物改变了 PFK-1 的构象，促进更多的酶分子转变为 R 态，从而提高了酶与 F-6-P 的亲和力，加快了反应速率。

在 F-2,6-BP 存在的情况下，PFK-1 与 ATP 的关系如图 22-20(2) 所示，图(2)的数据充分说明了 F-2,6-BP 的存在能够抵消高浓度 ATP 作为负别构效应物对 PFK-1 活性的抑制作用。



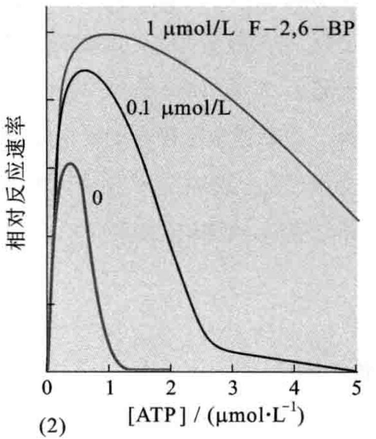


图 22-20 F-2,6-BP 对 PFK-1 活性的调节

四、丙酮酸激酶的调节

丙酮酸激酶"把守"着糖酵解的出口，因此也是一个很重要的调节位点。丙酮酸激酶活性有两种调节方式：一种是别构调节（图 22-21），另一种是共价修饰。前一种方式涉及的别构效应物有 ATP、丙氨酸和 F-1,6-BP。其中 ATP 和丙氨酸为负别构效应物，F-1,6-BP 为正别构效应物。ATP 抑制丙酮酸激酶活性的原理类似于 ATP 抑制 PFK-1，而 F-1,6-BP 以前馈激活（feed-forward activation）的方式作用于丙酮酸激酶，可保证通过 PFK-1 这一关的 F-1,6-BP 能够"走完"全程，还可防止其中的中间代谢物发生不必要的堆积。后一种方式是典型的"可逆的蛋白质磷酸化"。受胰高血糖素的作用，PKA 被激活，而被激活的 PKA 进一步催化丙酮酸激酶的磷酸化修饰，从而使其失活。

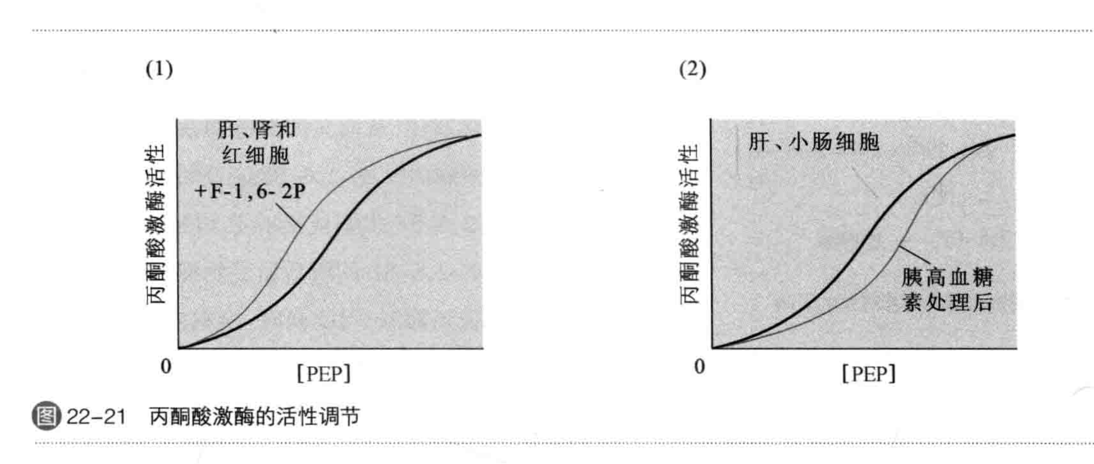


图 22-21 丙酮酸激酶的活性调节

框 22-2 理论联系实际——教你制作葡萄酒

如果你有兴趣，可以自己做葡萄酒，最好选择在暑假葡萄大量上市的时候酿制。

用料：葡萄 5 kg，白糖 1 kg，即 20% 左右（为一般发酵用的培养基含糖量）。

步骤：①将洗好的葡萄晾干，去梗。②用手把葡萄捏烂，皮肉分离即可。③取玻璃罐（5 L 罐，最好用小口，方便密封，不能用塑料瓶），一层葡萄一层白糖交叉地放入其中（注意白糖要放够，否则是酸的）。装瓶量达 4.5 L，留体积以备发酵时体积膨胀用。④全部放好后，将罐口密封，可以先用保鲜膜，再用多层塑料薄膜，要做到完全密封。⑤将密封好的罐子放入阴凉通风处，一般放置 40 d 左右。⑥如果用透明玻璃罐，可以很清楚地看到整个发酵过程，直到最后酒和残渣完全分离。这时罐口有些黏，而发酵产生的气体从里面渗漏使瓶口薄膜上突较高。⑦发酵完成后，可以用纱布将残渣和酒分离，只留下晶莹剔透的酒。5 kg 葡萄大约能酿出 3.75 kg 葡萄酒。

发酵注意事项：①要求合适的温度，理论温度为 20℃，而实际温度 15~25℃都可以，最好放在阴凉处。②不能用塑料瓶，因为塑料很可能会与酒精发生化学反应，产生一些有毒物质，危害人体健康。日常用陶瓷瓶、玻璃瓶均可以。③发酵结束前不要打开瓶盖，否则会改变发酵途径，产酸。④选用甜葡萄，否则单宁含量高，影响口感。⑤如果要做红葡萄酒就用紫葡萄，最好选择颜色乌黑的，这样酿出来的酒颜色才够红润。如果要做白葡萄酒就用绿葡萄。若要发酵更快，也可以在装罐时放入醪糟。⑥在酿酒的过程中，会出现白色泡沫，这是正常现象。⑦自己酿出来的酒一般度数不高，如果放上半年能达到 12°，跟市售的干红类似，味道也会更纯正一些。

e22-3 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/glycolysis.html（完全免费的医学生物化学课程网站，有关糖酵解的内容）
2. http://en.wikipedia.org/wiki/Glycolysis（维基百科有关糖酵解的内容）

3. http://www.science.smith.edu/departments/Biology/Bio231/glycolysis.html（美国史密斯学院提供的一段糖酵解途径的动画）
4. http://www.wiley.com/college/boyer/0470003790/animations/glycolysis/glycolysis.htm（糖酵解途径的动画，形象生动）
5. http://www.metabolicvisualizer.org/（内有糖酵解和糖异生途径的模拟显示）

参考文献：

1. Gancedo C, Flores C L. Moonlighting proteins in yeasts. *Microbiol Mol Biol Rev*. 72:197–210, 2008.
2. Hypoxia. HIF1 and glucose metabolism in the solid tumour. *Nat Rev Cancer*, 8:705–713, 2008.
3. Kim J, Dang C V. Multifaceted roles of glycolytic enzymes. *Trends Biochem Sci*, 30:142–150, 2005.
4. Dang C V, Semenza G L. Oncogenic alterations of metabolism. *Trends Biochem Sci*, 24:68–72, 1999.
5. Middleton R J. Hexokinases and glucokinases. *Biochem Soc Trans*, 18:180–183, 1990.
6. Hue L, Rider M H. Role of fructose-2,6-bisphosphate in the control of glycolysis in mammalian tissues. *Biochem J*, 245:313–324, 1987.

第二十三章　三羧酸循环

三羧酸循环（tricarboxylic cycle，TCA cycle）也称为柠檬酸循环（citric acid cycle）或 Krebs 循环，它发生在真核细胞的线粒体和原核细胞的细胞质基质中，绝对需要氧气的存在。作为一条不定向代谢途径，它既是糖类、脂质和蛋白质在细胞内最后氧化分解的共同代谢途径，又在很多生物分子的合成代谢中发挥重要的作用。

本章将重点介绍 TCA 循环的发现、反应步骤、生理功能以及调控机制。

第一节　三羧酸循环的发现

与 TCA 循环有关的研究最早始于 20 世纪初。当时有人发现，剁碎的动物组织悬液在无氧条件下，能够将某些小分子有机酸上的氢转移到一种叫亚甲蓝（又称甲烯蓝，methylene blue）的染料上，亚甲蓝因接受氢被还原，并从无色变为蓝色。这些有机酸包括琥珀酸（succinic acid）、延胡索酸（富马酸，fumaric acid）、苹果酸（malic acid）和柠檬酸（citric acid）等。但相同的动物组织悬液在有氧条件下，这些有机酸则被氧化成 CO_2 和 H_2O。

1937 年，Albert Szent-Györgyi 发现，在将上述少量的有机酸加到含有葡萄糖的组织悬液中以后，氧气的消耗量猛增，远远高于有机酸本身氧化所需要的量。于是 Albert Szent-Györgyi 认为，在有氧状态下，这些有机酸对葡萄糖的氧化具有催化作用。后来他又发现，琥珀酸脱氢酶的特异性抑制剂——丙二酸（malonic acid）能够抑制许多不同的动物组织悬液对氧气的利用。这就表明了琥珀酸是组织有氧代谢所必需的一种成分。

Quiz1 你认为 Krebs 为什么会选择鸽子的飞翔肌为研究对象？

基于上述结果，Hans Adolf Krebs 想搞清楚不同的有机酸在有氧代谢中的相互关系。他选用了鸽子的飞翔肌（flight muscle）作为实验材料，首先证实了 Albert Szent-Györgyi 的实验结果，同时还发现草酰乙酸（oxaloacetic acid）、α- 酮戊二酸（α-ketoglutaric acid）、异柠檬酸（isocitric acid）和顺乌头酸（*cis*-aconitic acid）也有类似的催化作用，以及这些有机酸对丙酮酸的氧化也具有促进作用。考虑到丙二酸抑制的是琥珀酸转变成延胡索酸的反应，Krebs 推测，琥珀酸脱氢酶催化的反应仅仅是丙酮酸完全氧化所涉及的全部反应中的一步。比较了所有对丙酮酸的氧化具有催化作用的有机酸之间的结构关系以后，Krebs 确定了它们之间进行生化转变的可能途径：柠檬酸→顺乌头酸→异柠檬酸→α- 酮戊二酸→琥珀酸→延胡索酸→苹果酸→草酰乙酸。但是，如果考虑到这些有机酸以催化剂量参与丙酮酸的氧化反应，机体内就必然存在某种再生机制，以保证它们能从一个单一的代谢中间物重新生成，而不会被完全消耗掉。显然，要是这些有机酸能组成一个环形的代谢通路，那么再生的问题也就迎刃而解了。可是，草酰乙酸如何变成柠檬酸呢？ Krebs 很快有了答案：在缺乏氧气的条件下，丙酮酸和草酰乙酸加入到肌肉组织悬液中可以形成柠檬酸。于是，他认为丙酮酸和草酰乙酸之间能够缩合成柠檬酸并释放 CO_2。

Krebs 提出的循环假说很快得到了其他一些实验证据的支持。证据之一是，在受到丙二酸抑制的肌肉组织切片中，加入环路中的任何有机酸（包括琥珀酸氧化的直接产物延胡索酸）均能导致琥珀酸的堆积；证据之二是，在大量的丙酮酸氧化受到丙二酸抑制的情况下，加入草酰乙酸、苹果酸或者延胡索酸都能刺激化学剂量的丙酮酸发生反应，但如果加入的是琥珀酸，就没有任何效果。

上述证据不仅证明了丙酮酸的氧化需要与草酰乙酸缩合成柠檬酸，还说明了随后的反应的确构

成了环式的代谢通路。Krebs 因此荣获了 1953 年的诺贝尔生理学或医学奖。

第二节　三羧酸循环的全部反应

在最初提出的 TCA 循环环路时，Krebs 并没有真正解释丙酮酸是如何与草酰乙酸缩合成柠檬酸的。实际上，丙酮酸并不能直接与草酰乙酸缩合成柠檬酸，它只有在转变成乙酰 -CoA 以后，由乙酰 -CoA 与草酰乙酸才能缩合成柠檬酸。关于丙酮酸转变成乙酰 -CoA 的反应已在第二十二章“糖酵解”中详细介绍过，但此反应并不是细胞内生成乙酰 -CoA 的唯一途径，在脂肪酸代谢和氨基酸代谢中也会生成乙酰 -CoA。

一、反应过程

完整的 TCA 循环共由 8 步反应组成，其中 3 步是不可逆反应(图 23-1)。

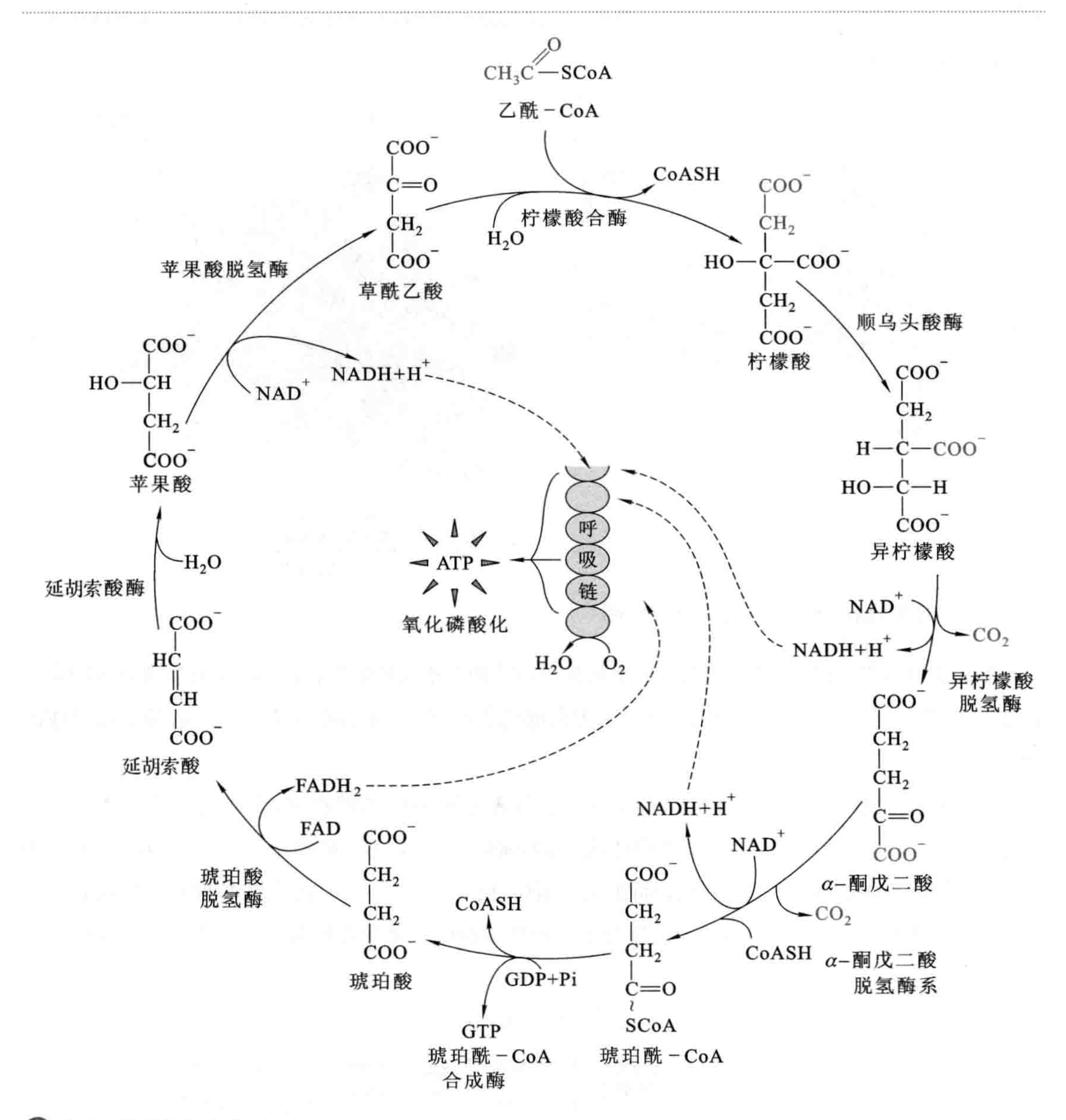


图 23-1　完整的 TCA 循环

1. 柠檬酸的合成

这是 TCA 循环的第一步反应，为不可逆反应，由柠檬酸合酶(citrate synthase)催化，反应的性质属

于有机化学里的羟醛缩合反应(aldol condensation)。

$$\underset{\text{乙酰-CoA}}{CH_3-C(=O)-S\text{-}CoA} + \underset{\text{草酰乙酸}}{O=\overset{2}{C}(-\overset{1}{C}OO^-)-CH_2-COO^-} \xrightarrow[\text{柠檬酸合酶}]{H_2O\ \ CoA\text{-}SH} \underset{\text{柠檬酸}}{HO-C(-CH_2-COO^-)(-COO^-)-CH_2-COO^-}\quad \Delta G^{\ominus\prime}=-32.2\ \text{kJ/mol}$$

$$\text{中间物：}\underset{\text{柠檬酰-CoA（反应中间物）}}{HO-C(-CH_2-C(=O)-S\text{-}CoA)(-COO^-)-CH_2-COO^-}$$

柠檬酸合酶通常由 2 个相同的亚基组成,它可被视为酶"诱导契合"模型的又一典型代表。如图 23-2 所示,在没有底物结合的时候,酶的 2 个亚基的构象是开放型(open form)的;一旦结合了底物,则被诱导为紧密型或封闭型(closed form)。其反应动力学为序列有序(sequential ordered)型,在反应中,草酰乙酸首先与酶活性中心结合,这种结合迅速诱导活性中心的构象发生显著的变化,从而创造出乙酰 -CoA 的结合位点。随后,乙酰 -CoA 结合到酶的活性中心,并与草酰乙酸形成中间产物——柠檬酰 -CoA,这时酶的构象再次发生变化,远离活性中心的一个关键的 Asp 残基被拉到柠檬酰 -CoA 的硫酯键附近,硫酯键很快被切开,终产物辅酶 A 和柠檬酸被依次释放。

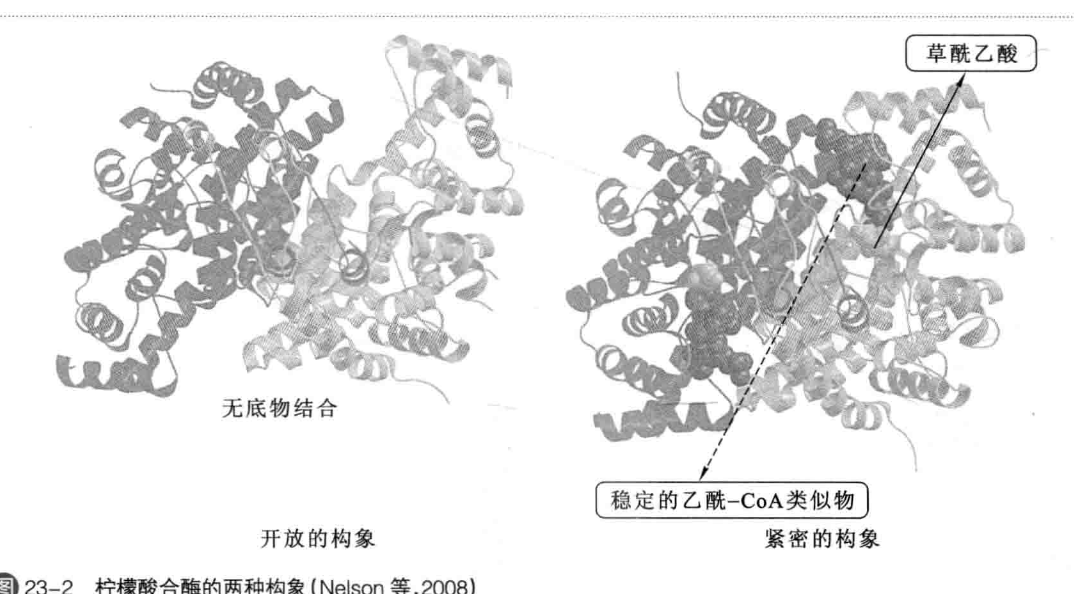


图 23-2 柠檬酸合酶的两种构象(Nelson 等,2008)

在催化过程中,柠檬酸合酶所发生的由底物草酰乙酸和中间产物柠檬酰 -CoA 分别诱导的两次构象变化,既防止了乙酰 -CoA 的提前释放,又大大降低了乙酰 -CoA 在活性中心被 Asp 残基水解成乙酸的可能性。

Quiz2 糖酵解中有哪一种抑制剂的作用方式与氟代乙酸相似?

乙酸的类似物——氟代乙酸(fluroacetate)是 TCA 循环强烈的抑制剂,因此有人将其作为灭鼠药的成分。它也存在于一些有毒的植物体内,例如疯草(locoweed)。动物吃了这一类植物会发生中毒。然而氟代乙酸在细胞内并不直接抑制 TCA 循环,而是先在乙酰 -CoA 合成酶催化下形成氟代乙酰 -CoA,氟代乙酰 -CoA 再与草酰乙酸缩合生成氟代柠檬酸。氟代柠檬酸是催化下一步反应的酶即

$$\underset{\text{柠檬酸}}{HO-C(-CH_2COO^-)(-COO^-)-CH(H)-COO^-} \underset{\text{顺乌头酸酶}}{\overset{H_2O}{\rightleftharpoons}} \underset{\text{顺乌头酸}}{\left[C(-CH_2COO^-)(-COO^-)=C(H)-COO^-\right]} \underset{\text{顺乌头酸酶}}{\overset{H_2O}{\rightleftharpoons}} \underset{\text{异柠檬酸}}{H-C(-CH_2COO^-)(-COO^-)-C(H)(OH)-COO^-}\quad \Delta G^{\ominus\prime}=13.3\ \text{kJ/mol}$$

图 23-3 氟代乙酸在细胞内的代谢转变及其对 TCA 循环的影响

顺乌头酸酶的抑制剂(图 23-3)。

2. 异柠檬酸的形成

这是一步将柠檬酸转变为异柠檬酸的异构化反应。反应经过顺乌头酸中间物,由顺乌头酸酶(*cis*-aconitase)催化。

顺乌头酸酶本质上是一个铁硫蛋白(Fe_4S_4,参看第二十章“生物氧化”),其分子之中的铁硫簇是催化反应所必需的,其中的一个铁作为路易斯酸参与催化(图 23-4)。

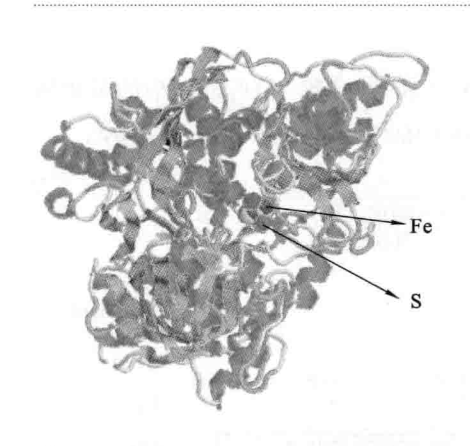


图 23-4 铁硫蛋白在顺乌头酸酶反应中的作用(Nelson 等,2008)

已在哺乳动物细胞内发现两种形式的顺乌头酸酶:一种位于线粒体基质参与 TCA 循环;另外一种位于细胞质基质,其功能主要是作为细胞内铁离子的感应器,参与铁蛋白与转铁蛋白受体在翻译水平上的表达调控(详见第四十三章“真核生物的基因表达调控”)。

需要特别注意的是,在形成的异柠檬酸分子中,羟基只会与来自草酰乙酸而绝对不会与来自乙酰 -CoA 的β- 碳原子相连(图 23-5)。这里涉及的机制可以用酶与底物结合的“三点附着”模型来解释(参见第八章“酶学概论”)。

尽管在反应达到平衡时(在 25℃和 pH 7.4 条件下),柠檬酸、顺乌头酸和异柠檬酸各占 90%、4% 和 6%,但在下一步高度放能的不可逆反应中,异柠檬酸作为底物很容易被消耗,致使这一步反应仍然能以较快的速率向正反应方向进行。

图 23-5 顺乌头酸酶催化反应中产物的立体专一性

柠檬酸转变成异柠檬酸是 TCA 循环必不可少的一步,原因是柠檬酸在结构上并非氧化反应的良好底物,但异柠檬酸却不一样,经过异构化,其三级羟基变成了易氧化的二级羟基。

3. 异柠檬酸的氧化脱羧

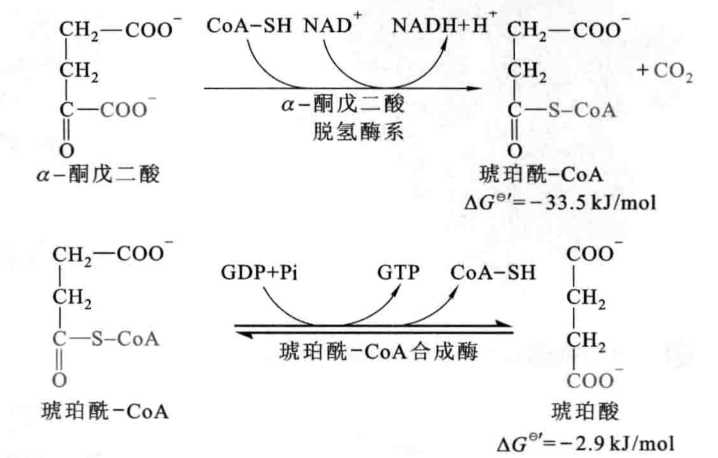
这是一步不可逆的氧化脱羧反应,由异柠檬酸脱氢酶催化,反应分为两步:先是脱氢,形成草酰琥珀酸(oxalosuccinate);然后是β- 脱羧,产生 CO_2 和α- 酮戊二酸。

异柠檬酸 —(异柠檬酸脱氢酶，脱氢；$NAD(P)^+$ → $NAD(P)H + H^+$)→ [草酰琥珀酸] —(β-脱羧，$-CO_2$)→ α-酮戊二酸

已发现两种形式的异柠檬酸脱氢酶，分别使用辅酶Ⅰ和辅酶Ⅱ作为氢的受体。但一般认为，参与TCA循环的主要是辅酶Ⅰ。如果辅酶Ⅰ作为氢的受体，形成的NADH就可以直接从线粒体内膜上的复合体Ⅰ进入呼吸链，通过氧化磷酸化产生ATP；如果是辅酶Ⅱ参与反应，形成的NADPH就可以作为还原剂参与生物合成。

4. α-酮戊二酸的氧化脱羧

这一步反应与丙酮酸的氧化脱羧十分相似(参看第二十二章“糖酵解”中的丙酮酸的氧化脱羧)，由α-酮戊二酸脱氢酶、二氢硫辛酸转琥珀酰酶(dihydrolipoyl transsuccinylase)和二氢硫辛酸脱氢酶3个酶组成的α-酮戊二酸脱氢酶系催化。从总反应式看出，抑制丙酮酸氧化脱羧的亚砷酸也能强烈抑制这一步反应。



5. 底物水平的磷酸化

这是TCA循环内唯一的一步底物水平磷酸化反应，由琥珀酰-CoA合成酶(succinyl-CoA synthetase)催化，酶的命名根据逆反应的性质而定。

琥珀酰-CoA合成酶也被称为琥珀酸硫激酶(succinate thiol kinase)，其催化过程牵涉到一系列高能分子的形成和转变，因此能量的损失微乎其微，$\Delta G^{\ominus\prime}=-2.9$ kJ/mol。

长期以来，人们曾一直认为，动物组织中的琥珀酰-CoA合成酶只能催化GDP的磷酸化，植物和微生物的琥珀酰-CoA合成酶主要催化ADP的磷酸化，但后来发现，某些动物组织同时有分别对GDP和ADP专一性的两种琥珀酰-CoA合成酶。

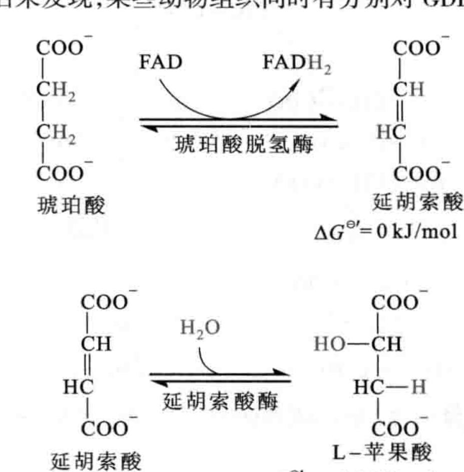
6. 琥珀酸的脱氢

这是TCA循环中的第三次脱氢反应，由琥珀酸脱氢酶催化，产物是反丁烯二酸。反丁烯二酸也被称为延胡索酸或富马酸。琥珀酸的类似物丙二酸是该酶的竞争性抑制剂。

琥珀酸脱氢酶是复合体Ⅱ的主要成分，它的结构与功能参看第二十章“生物氧化”。

7. 苹果酸的形成

这是一步水合反应，由延胡索酸酶催化。



8. 草酰乙酸的再生

这是TCA循环的最后一步反应，也是TCA循环中的第四次氧化还原反应，由苹果酸脱氢酶催化。尽管在热力学上极不利于正反应的进行，可是在体内，反应产物草酰乙酸可以迅速被下一步不可逆反应所消耗，NADH则进入呼吸链被彻底氧化，因此整个反应被“强行拉向”正反应。

$$\text{L-苹果酸} + NAD^+ \xrightarrow[\text{苹果酸脱氢酶}]{} \text{草酰乙酸} + NADH + H^+$$

$\Delta G^{\ominus\prime} = 29.7\ \text{kJ/mol}$

二、三羧酸循环小结

TCA 循环的总反应式为：

$$\text{乙酰-CoA}+3NAD^++FAD+GDP+Pi+2H_2O \longrightarrow 2CO_2+3NADH+FADH_2+GTP+2H^++CoA$$

其中几个重要的特征归纳如下：

(1) 有 2 个 C 原子进入循环，但在随后的反应中又有 2 个 C 原子离开。需要特别注意的是，离开的 2 个 C 原子并非是最初进入的那 2 个，而是来源于草酰乙酸。

Quiz4 如果进入首轮 TCA 循环的乙酰 -CoA 上的两个 C 用 ^{14}C 标记，那么在随后的循环反应中，带有 ^{14}C 标记的 CO_2 何时产生？

(2) 有 4 对氢原子离开循环，并进入呼吸链进一步氧化产生更多的 ATP。

(3) 有一步底物水平磷酸化反应，产生的能量货币是 GTP 或 ATP。

(4) 有 2 个水分子被消耗，分别作为柠檬酸合酶和延胡索酸酶的底物。

(5) 氧气并不直接参与循环，但它的存在是 TCA 循环正常进行所必需的。只有在有氧的条件下，NAD^+ 和 FAD 才可以通过呼吸链的氧化而得以再生，循环内的 4 步氧化还原反应才可以持续进行，否则循环最终会因为 NAD^+ 和 FAD 的缺乏而停止。这一点与糖酵解不一样，糖酵解在无氧或有氧的条件下都可以进行，是因为在两种条件下 NAD^+ 均可再生。

第三节　三羧酸循环的生理功能

TCA 循环作为需氧生物最重要的代谢途径之一，其主要的生理功能包括：

(1) 作为需氧生物细胞内所有代谢燃料最终氧化分解的共同代谢途径。

(2) 提供多种生物分子合成的前体，参与合成代谢(图 23-6)。

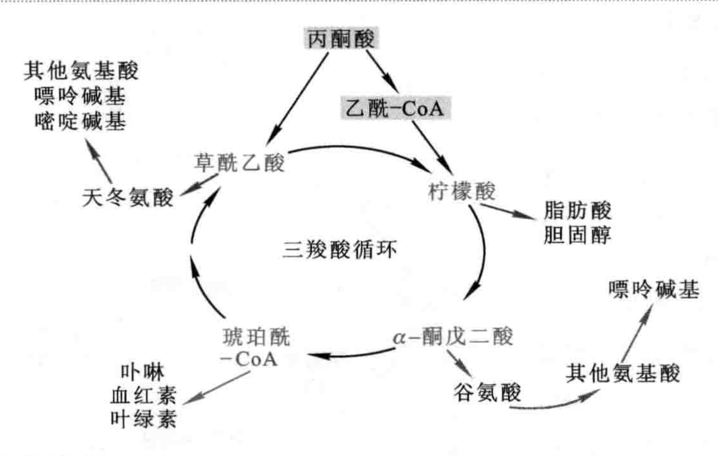


图 23-6　TCA 循环中间物的去向

(3) 与呼吸链偶联可产生更多的 ATP　1 分子葡萄糖彻底氧化成 CO_2 和 H_2O 能够产生 30~32 分子的 ATP，其计算过程参考表 23-1，其中由糖酵解所产生的 ATP 所占的比例并不多，其终产物丙酮酸经丙酮酸脱氢酶系，再经 TCA 循环和氧化磷酸化可以产生更多的 ATP。其中 2 分子乙酰 -CoA 经 TCA 循环彻底氧化产生的 ATP 分子数目应该是 20，远远高于糖酵解途径。

Quiz5 如果 TCA 循环的琥珀酸脱氢酶反应受到丙二酸的完全抑制，那么 1 分子葡萄糖在肝细胞内氧化、分解最多可产生多少 ATP？

表 23-1　1 分子葡萄糖彻底氧化过程中的 ATP 收支情况

与 ATP 合成相关的反应	合成 ATP 的方式	合成 ATP 的量
糖酵解（包括氧化磷酸化）		5 或 6 或 7
己糖激酶	消耗 ATP	-1
PFK-1	消耗 ATP	-1
磷酸甘油酸激酶	底物水平磷酸化	2
丙酮酸激酶	底物水平磷酸化	2
3- 磷酸甘油醛脱氢酶（NADH）	氧化磷酸化	3、4 或 5（取决于 NADH 通过何种途径进入呼吸链）
丙酮酸脱氢酶系	氧化磷酸化	2 × 2.5=5
TCA 循环		20
异柠檬酸脱氢酶（NADH）	氧化磷酸化	2.5 × 2=5
α- 酮戊二酸脱氢酶系（NADH）	氧化磷酸化	2.5 × 2=5
琥珀酰 -CoA 合成酶	底物水平磷酸化	1 × 2=2
琥珀酸脱氢酶（$FADH_2$）	氧化磷酸化	1.5 × 2=3
苹果酸脱氢酶（NADH）	氧化磷酸化	2.5 × 2=5
总 ATP 量		30、31 或 32

(4) 循环中的某些中间产物可作为别构效应物，去调节其他代谢途径　例如，柠檬酸可作为 PFK-1 的负别构效应物和 1，6- 二磷酸果糖磷酸酶的正别构效应物，分别调节糖酵解和糖异生（参看第二十二章“糖酵解”和第二十五章“糖异生”）。此外，柠檬酸还可以作为正别构效应物，去刺激脂肪酸合成的限速酶——乙酰 -CoA 羧化酶的活性（参看第二十九章“脂肪酸代谢”）。

(5) 产生 CO_2　1 分子乙酰 -CoA 进入 TCA 循环，最后有 2 个碳变成了 CO_2。CO_2 的功能往往容易被忽视，实际上它在生物体内除了能够调节酸碱平衡以外，还可以作为羧化反应碳单位的供体。

第四节　乙醛酸循环

植物、某些微生物（如大肠杆菌和酵母）以及无脊椎动物能够使用乙酸或者其他一些能够在细胞内转化为乙酰 -CoA 的物质，来作为唯一的碳源和生物合成的前体。这是因为在它们的体内存在一种特殊的代谢途径，即乙醛酸循环（glyoxylate cycle）。该途径实际上是 TCA 循环的变化形式。

乙酸并不能直接进入乙醛酸循环，它在细胞内必须首先被活化成乙酰 -CoA。乙醛酸循环所涉及的反应以及它与 TCA 循环的比较如图 23-7 所示，它与 TCA 循环的主要差别包括：①在每一轮循环中，前者有 2 分子乙酰 -CoA 进入，后者只有 1 分子乙酰 -CoA 进入；②乙醛酸循环只产生 NADH，不产生 $FADH_2$；③乙醛酸循环无底物水平磷酸化反应，因此无法通过此途径合成 ATP；④乙醛酸循环不生成 CO_2，但每一轮 TCA 循环有 2 分子 CO_2 的释放。

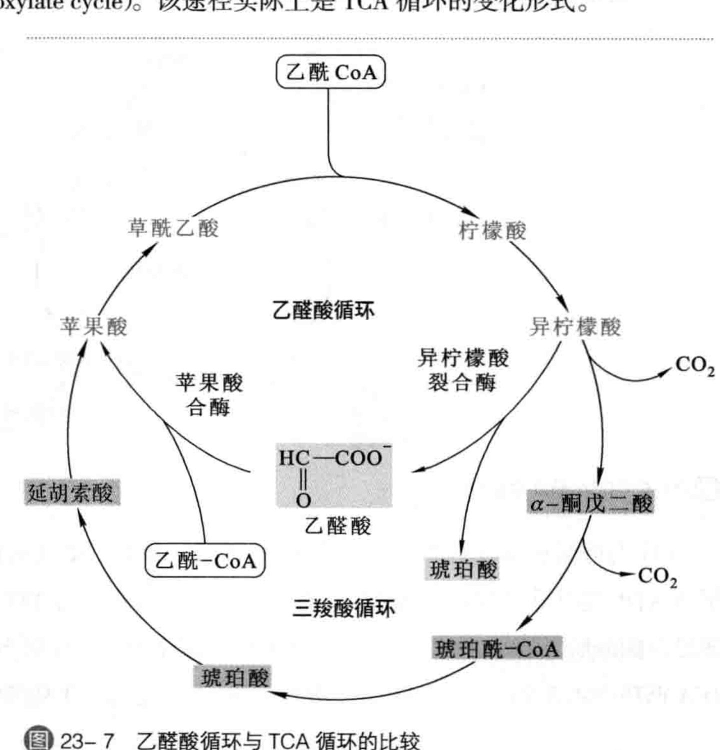


图 23-7　乙醛酸循环与 TCA 循环的比较

乙醛酸循环与 TCA 循环在性质上的差异说明了两者在功能上一定有所不同。对于同时含有这两

种代谢途径的生物来说，异柠檬酸实际上是一个“分水岭”。细胞内的异柠檬酸是在TCA循环中被异柠檬酸脱氢酶催化形成α-酮戊二酸，还是在乙醛酸循环中被异柠檬酸裂合酶（isocitrate lyase）裂解成乙醛酸和琥珀酸，完全取决于细胞当时的能量状态。

细胞能量状态既能够影响到异柠檬酸裂合酶的活性，又能影响到一种特殊的蛋白质激酶的活性，而该蛋白质激酶同时具有磷酸酶的活性，因此是一种双功能酶（与前面遇到的PFK-2相似）。该双功能酶通过磷酸化和去磷酸化的手段直接控制异柠檬酸脱氢酶的活性（磷酸化的异柠檬酸脱氢酶没有活性，去磷酸化的异柠檬酸脱氢酶具有活性）。当细胞的能量状态较低的时候，高水平的ADP和AMP，以及糖酵解和TCA循环的某些中间产物作为别构效应物，一方面抑制异柠檬酸裂合酶的活性，另一方面抑制上述蛋白质激酶的活性，因此这时候的异柠檬酸脱氢酶以具有活性的去磷酸化形式存在，TCA循环得以正常地进行，并产生足够的ATP；相反，在细胞的能量状态较高的时候，上述的蛋白质激酶的活性被激活，而磷酸酶的活性被抑制，异柠檬酸脱氢酶被磷酸化成无活性的形式，同时，异柠檬酸裂合酶活性被激活，结果便是TCA循环受到抑制，乙醛酸循环被激活。

乙醛酸循环的生理意义在于：它使具有此途径的生物能够以乙酸作为唯一的碳源。每当1分子乙酰-CoA进入TCA循环时，最后总有2个碳原子以CO_2的形式被丢失。然而，乙酰-CoA进入乙醛酸循环，就跳过了两次脱羧反应，并无碳单位的损失，而是净合成了糖异生的前体——苹果酸。苹果酸可离开乙醛酸循环体进入细胞质，再经苹果酸脱氢酶的催化，转变为草酰乙酸，草酰乙酸可以作为糖异生的原料（参看第二十五章“糖异生”）。琥珀酸则离开乙醛酸循环体进入线粒体，经TCA循环形成草酰乙酸，然后在谷丙转氨酶的催化下，接受谷氨酸上的氨基转变为天冬氨酸。天冬氨酸离开线粒体进入乙醛酸循环体，再重新转变为草酰乙酸，以维持草酰乙酸的再生。

在植物细胞中，乙醛酸循环被限制在一种特定的细胞器即乙醛酸循环体（glyoxysome）内，但这种细胞器并不出现在植物所有的细胞，也不会始终存在。当富含油脂的种子发芽的时候，乙醛酸循环体即在细胞内形成。形成的乙醛酸循环体使得脂肪在植物细胞内转变成葡萄糖成为可能，这对于尚不能通过光合作用合成葡萄糖的细胞来说尤为重要。

第五节　三羧酸循环的回补反应

在理论上，虽然只要有1分子草酰乙酸就可以让TCA循环不断地进行下去，但如果草酰乙酸太少，TCA循环只会以极低的速率进行。由于TCA循环中的多种代谢中间产物可被其他代谢途径利用作为生物合成的原料，因此细胞内若没有某些机制及时补充被“挪作他用”的草酰乙酸或其他代谢中间产物的话，TCA循环迟早就会受到影响。实际上，细胞已经预备了若干回补反应（anaplurotic reaction），可以及时补充被其他途径消耗的TCA循环的中间产物，提高细胞对“二碳单位”的氧化速率。

TCA循环的回补反应主要在草酰乙酸、α-酮戊二酸、琥珀酰-CoA和苹果酸这4个位点进行，现分述如下：

（1）草酰乙酸的回补　这是回补反应的主要形式，相关的酶有PEP羧化酶（PEP carboxylase）、丙酮酸羧化酶（pyruvate carboxylase）和PEP羧激酶（PEP carboxykinase）（图23-8）。这3种酶在不同生物和不同组织中的分布并不相同，其中PEP羧化酶主要分布在细菌、酵母和高等植物中，丙酮酸羧化酶主要存在于动物的肝和肾，而PEP羧激酶则主要分布在动物的心肌和骨骼肌。

（2）α-酮戊二酸的回补　由谷丙转氨酶催化的转氨基反应，或者谷氨酸脱氢酶催化的氧化脱氨基反应均可以将谷氨酸转化为α-酮戊二酸，且都发生在线粒体基质，因此都能充当回补反应。

（3）琥珀酰-CoA的回补　Ile、Val、Met和Thr这4种氨基酸在细胞内均可以被氧化成琥珀酰-CoA。此外，奇数脂肪酸在细胞内氧化也可以产生琥珀酰-CoA（参看第二十九章“脂肪酸代谢”），这些产生

图 23-8 草酰乙酸的回补反应

琥珀酰 -CoA 的反应均可以作为回补反应使用。

Quiz6 有一则广告，声称苹果酸能够减肥。你认为这个广告可信吗？

(4) 苹果酸的回补 这是由苹果酸酶(malic enzyme)催化的回补反应，此回补反应广泛存在于各种真核生物和原核生物之中。

第六节 三羧酸循环的调控

为了适应细胞对能量的需求，细胞内的 TCA 循环是受到严格控制的。对 TCA 循环本身的调控集中在柠檬酸合酶、异柠檬酸脱氢酶和 α- 酮戊二酸脱氢酶(图 23-9)，其中最重要的调控位点是异柠檬酸脱氢酶，其次是 α- 酮戊二酸脱氢酶，至于柠檬酸合酶，对它的调控多见于原核生物。

乙酰 -CoA 是 TCA 循环第一步反应的底物，因此机体还可以通过控制它的形成来控制 TCA 循环，而这必然牵涉到细胞内参与乙酰 -CoA 合成的酶，例如丙酮酸脱氢酶系和参与脂肪酸 β- 氧化的酶。

1. 柠檬酸合酶的调控

对柠檬酸合酶的调控方式主要为别构调节。细胞高能状态的指示剂、反应的中间产物或终产物(ATP、NADH 和琥珀酰 -CoA)均可作为负别构效应物来抑制该酶的活性，而细胞低能状态的指示剂 ADP 可作为正别构效应物来刺激该酶的活性。此外，高浓度的柠檬酸可以通过竞争的方式反馈抑制柠檬酸合酶。

2. 异柠檬酸脱氢酶的调控

对异柠檬酸脱氢酶的调控有两种方式：第一种是植物细胞才有的共价修饰；第二种是别构调节。其中第一种方式已在乙醛酸循环中做过专门的介绍，这里只讨论第二种方式。作为异柠檬酸脱氢酶负别构效应物的是 ATP 和 NADH，正别构效应物是 ADP 和 Ca^{2+}。细胞选择 Ca^{2+} 作为异柠檬酸脱氢酶以及后面将要讨论的 α- 酮戊二酸脱氢酶和丙酮酸脱氢酶的正别构效应物，原因是 Ca^{2+} 为肌肉收缩的信号，实际也就是肌细胞需要 ATP 的信号。

3. α- 酮戊二酸脱氢酶系的调控

对 α- 酮戊二酸脱氢酶的调控与对丙酮酸脱氢酶系的调控方法相似，但没有共价修饰的形式，只有别构调控和产物的竞争性反馈抑制。其中，Ca^{2+} 和 AMP 别构激活 α- 酮戊二酸脱氢酶，琥珀酰 -CoA

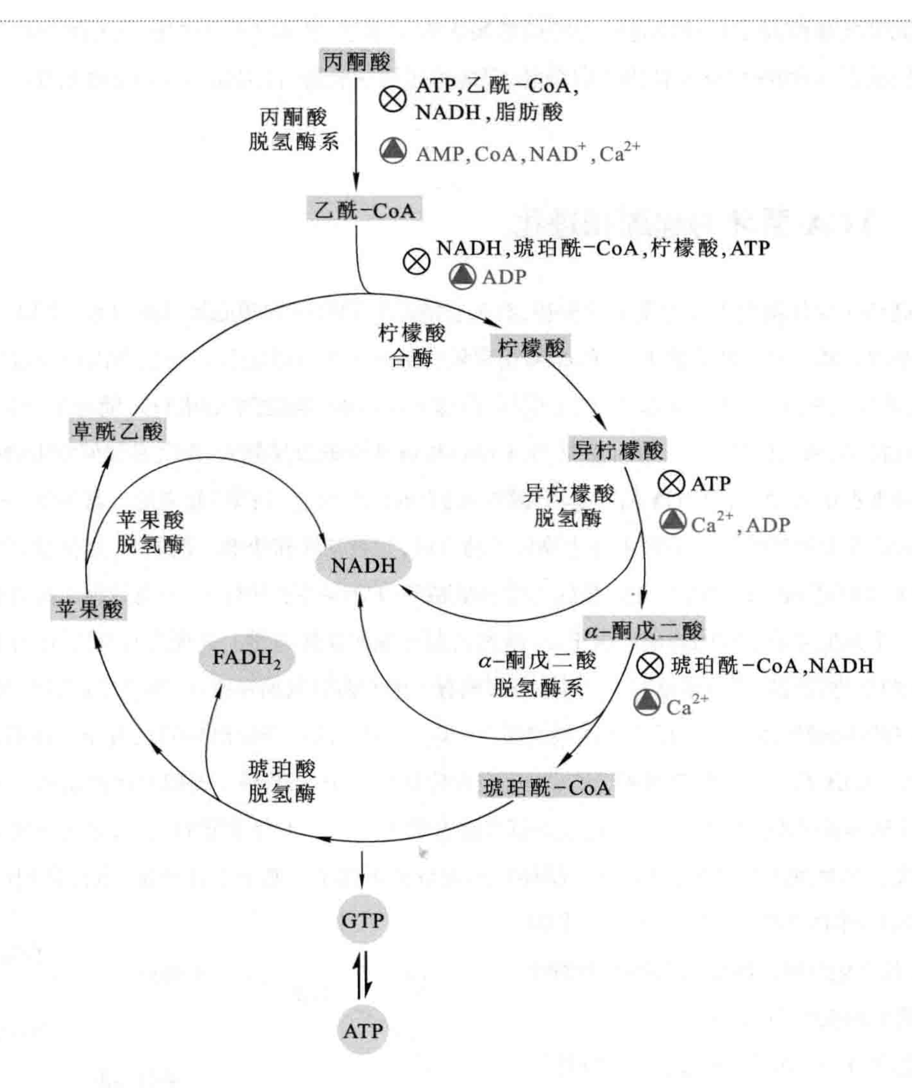


图 23-9　TCA 循环途径及其调控（Nelson 等，2008）

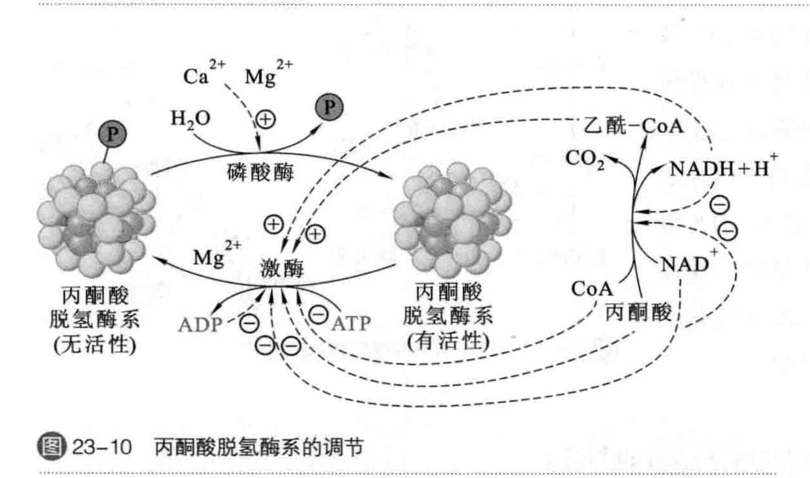


图 23-10　丙酮酸脱氢酶系的调节

和 NADH 分别反馈抑制二氢硫辛酸转琥珀酰酶和二氢硫辛酸脱氢酶。

4. 丙酮酸脱氢酶系的调控

对丙酮酸脱氢酶系的调控有产物的竞争性反馈抑制、别构调节和丙酮酸脱氢酶的共价修饰 3 种形式，最后一种仅存在于真核生物（图 23-10）。乙酰 -CoA 和 NADH 分别反馈抑制二氢硫辛酸转乙酰酶（E_2）和二氢硫辛酸脱氢酶（E_3）。参与丙酮酸脱氢酶别构调节的效应物仍然主要是细胞能量状态的指示剂，如 ADP、NAD^+ 和 NADH。

参与丙酮酸脱氢酶共价修饰的酶是丙酮酸脱氢酶激酶和磷酸酶，两者都紧密结合于丙酮酸脱氢酶系，其中激酶催化丙酮酸脱氢酶在特定位点的丝氨酸残基的磷酸化，从而导致其活性丧失，磷酸酶的功能正好相反。这两种酶的活性又受到别构效应物的控制，其中 ADP、CoA、NAD^+ 和丙酮酸能够抑制激酶的活性，Ca^{2+} 能刺激磷酸酶的活性，这 5 种物质共同维持丙酮酸脱氢酶处于有活性的去磷酸化状态，而乙酰 -CoA 和 NADH 作为正别构效应物激活激酶的活性，有利于丙酮酸脱氢酶主要以无活性的磷酸化形式存在。

Quiz7　双氯乙酸是丙酮酸脱氢酶激酶的一种特异性抑制剂，如果就将它加入到肿瘤组织之中，可导致癌细胞的大量死亡，对此你如何解释？

某些高乳酸血症患者的病因是体内的磷酸酶丧失了活性，从而导致丙酮酸脱氢酶始终处于无活性的磷酸化状态，丙酮酸因此在体内难以分解，只能被还原成乳酸，最终便出现高乳酸血症。

第七节　TCA 循环的起源和进化

TCA 循环作为代谢世界极为重要的枢纽，有关它的起源和最初的功能始终能引起人们无限的遐想和十分浓厚的兴趣。它与糖酵解不一样，必须在有氧的条件下才可以进行。在生命出现和进化的早期阶段，大气中缺乏氧气，因此不可能有 TCA 循环，但这并不影响到糖酵解的进行。随着生命的进化，出现了光合有机体，氧气随之产生，可以想象 TCA 循环也就渐渐地发展起来，但究竟如何发生的呢？

有一种观点认为，最初的 TCA 循环是由细胞中已有的化学反应"拼凑"起来的。其中的一个证据来自循环中的许多中间产物参与多种重要生物分子的合成，如氨基酸和卟啉。很显然，这些重要生物分子合成的前体，如丙酮酸、α-酮戊二酸、草酰乙酸和琥珀酸等应该很早就有了，否则细胞无法合成氨基酸和卟啉等对于其生存所必需的物质。由于 α-酮酸的氧化脱羧在热力学上是极为有利的，因此有理由认为 α-酮酸的氧化脱羧反应很早就有了。只是当细胞有了丙酮酸脱氢酶系或 α-酮戊二酸脱氢酶系以后，脱羧、氧化和硫酯键形成这三个反应才能被偶联在一起，而脱羧反应释放出来的自由能正好用来驱动乙酰-CoA 和 NADH 的合成。显而易见，这三步反应也应该早于 TCA 循环。支持上面假说的另外一个证据是由谷草转氨酶催化的反应。此反应也是氨基酸合成中的一个十分重要的反应，它可直接导致草酰乙酸的形成。当草酰乙酸产生以后，TCA 循环的出现也就不远了。基于上述证据，也许最初的 TCA 循环所包括的代谢中间产物的数目少于后来的 TCA 循环，其氧化分解丙酮酸或其他物质的效率也远远低于现在的 TCA 循环。

在原始的 TCA 循环出现以后，它当时的功能究竟是什么呢？当我们从逆时针方向去观察 TCA 循环的时候，似乎有了答案：CO_2 被同化了（图 23-11）。这是不是就是 TCA 循环最初的功能？如果是，又该如何驱动这个十分耗能的反应？对于这个问题，有人认为是来自 H_2S 与 FeS 反应释放出的能量。H_2S 与 FeS 反应的产物为 FeS_2，已发现这种产物在古老的地质年代中含量极为丰富，但它是否就是原始的铁硫中心还有待于研究。

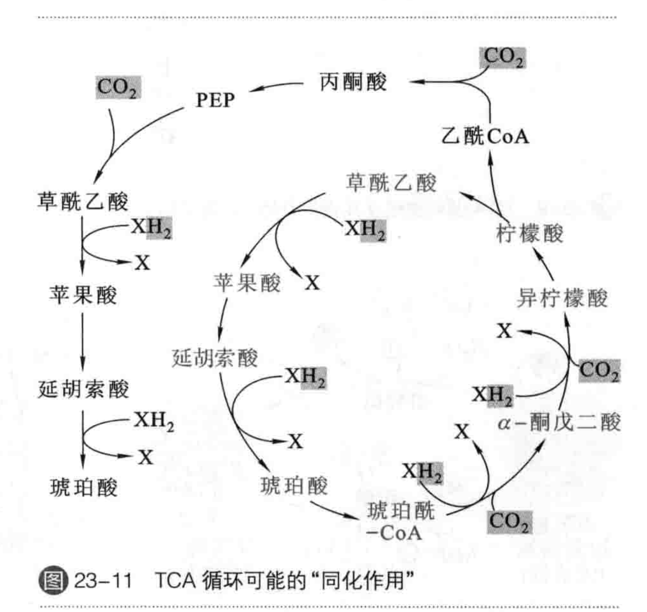


图 23-11　TCA 循环可能的"同化作用"

框 23-1　理论联系实际——短跑、中跑与长跑的能量消耗

奥运会 100 m 短跑比赛 10 s 左右就基本结束了，而马拉松比赛却要两个多小时才能结束。这两种运动都是赛跑，但它们利用的能源却有所不同。

先看看短跑比赛。短跑运动员能量的直接来源是存储在 ATP 分子中的，但储存的 ATP 数量非常有限，只有约 85 g。这点能量只能让短跑运动员跑约 3 s。显然，机体还要动用另一个能量储存——磷酸肌酸。磷酸肌酸也储存在肌肉中，在需要的时候，它能将其储存的高能磷酸键转移给 ADP，以补充新的 ATP。这些磷酸肌酸贮备的能量还可让运动员再跑五、六秒，这基本上够一个运动员跑完全程，等比赛结束的时候，运动员体内的磷酸肌酸已被全部耗尽。但在现实中，所有的运动员在冲刺阶段的最后都在相对放缓，这是因为能源正在枯竭，最后的冠军往往是在最后冲刺阶段减速最少的人！

再看看中跑比赛—— 400 m或800 m赛跑。这些选手需要跑更长的时间。当他们消耗所有的ATP和磷酸肌酸贮备以后，就依赖于细胞合成新的ATP了。实际上，细胞一直在制造ATP，但是因为跑步的时候对ATP的需求量很大，细胞的产能反应便加快了。然而选手们运动得太剧烈了，导致氧气的供应跟不上，于是肌肉不得不更依赖于糖酵解和乳酸发酵提供能量。在这样的情况下，乳酸在肌肉中积累，并扩散进入血流。积累在肌肉中的乳酸会不可避免地导致肌肉疲劳，最终可导致肌肉衰竭，因此这样的运动不能持续很久。

最后看看马拉松比赛，它呈现的是一个完全不同的场景。选手就像短跑和中跑选手一样，很快就消耗了贮备的ATP和磷酸肌酸。随后，糖酵解、三羧酸循环和氧化磷酸化开始提速。与中、短跑选手在ATP生成上的主要差别是：长跑选手的肌肉几乎全部利用有氧代谢产生的ATP。他们赛跑的速率可以使他们的氧气供应没有问题。实际上，只有有氧代谢才能提供连续数小时的运动。理论上，这样的条件可以使选手先动用其贮备的糖原，然后是脂肪，从而可以一直跑下去，但现实的因素使选手不可能一直跑下去，如脱水和疲劳的影响。由此可见，长跑运动员需要极佳的利用线粒体有氧代谢产生ATP的能力，而中、短跑运动员需要更强地利用细胞质基质的无氧代谢产生ATP的能力。

有趣的是，不同类型的运动员在肌肉结构上的差别反映了他们不同的需求。肌肉组织一般由两种主要的肌纤维组成：一种是快肌，大而丰满，只有少量的线粒体，呈白色，贮备了大量的糖原并含有高水平的糖酵解和乳酸发酵所需的酶。这种肌纤维很快就会疲劳，原因是发酵产能的效率低。另一种是慢肌，其细胞的直径只有快肌的一半，呈红色，内有大量的肌红蛋白，可贮存更多氧气。此外，其内还含有大量的线粒体。由于含有丰富的氧气和大量的线粒体，因此慢肌更适合进行有氧呼吸，这对于长跑运动员来说是有益的。已有研究发现，短跑运动员的肌肉中含有更多的快肌，而需要耐力的长跑运动员则有更多的慢肌纤维。但一个人肌肉中快肌和慢肌的相对含量是由遗传因素决定的，还是受到后天训练的影响？对这个问题的回答是：两个方面都有作用。如果单纯考虑遗传因素的话，也许某些人天生适合做短跑运动，像当今100 m跑世界纪录的保持者博尔特，其肌肉中存在更多的快肌纤维；相反，某些人更适合做马拉松选手。

推荐网址：

e23-1 本章小结填填看

1. http://en.wikipedia.org/wiki/Citric_acid_cycle（维基百科有关三羧酸循环的内容）
2. http://themedicalbiochemistrypage.org/tca-cycle.html（完全免费的医学生物化学课程网站，有关三羧酸循环的内容）
3. http://en.wikibooks.org/wiki/Principles_of_Biochemistry/Krebs_cycle_or_Citric_acid_cycle（维基书库提供的生物化学原理，有关三羧酸循环的内容）
4. http://www.wiley.com/college/boyer/0470003790/animations/tca/tca.swf（内有三羧酸循环的动画，形象、直观）
5. http://web.virginia.edu/Heidi/chapter20/chp20.htm（美国弗吉尼亚大学生物化学网络课程，有关三羧酸循环的内容）

参考文献：

1. Eastmond P J, Graham I A. Re-examining the role of the glyoxylate cycle in oilseeds. *Trends Plant Sci*, 6: 72–77, 2001.
2. Johnson J D, et al. Genetic evidence for the expression of ATP- and GTP-specific succinyl-CoA synthetases in multicellular eucaryotes. *J Biol Chem*, 273(42): 27 580–27 586, 1998.
3. Wolodko W T, et al. The crystal structure of succinyl-CoA synthetase from Escherichia coli at 2.5-Å resolution. *J Biol Chem*, 269: 10 883–10 890, 1994.
4. Weigand G, Remington S J. Citrate synthase: structure, control, and mechanism. *Annu Rev Biophys Biophys Chem*, 15: 97–117, 1986.

第二十四章　磷酸戊糖途径

磷酸戊糖途径(phosphate pentose pathway, PPP)又名磷酸己糖支路(hexose monophosphate shunt, HMS)或6–磷酸葡糖酸途径(6–phosphogluconate pathway)。它发生在细胞质基质中,是糖代谢的又一条重要途径。虽然此条代谢途径并不产生ATP和NADH,但能够产生NADPH和核糖这两种十分重要的生物分子,因此它的重要性并不亚于糖酵解和三羧酸循环。

本章将重点介绍磷酸戊糖途径的反应、功能和调控。

第一节　磷酸戊糖途径的全部反应

一条完整的磷酸戊糖途径由8步反应构成,根据反应的性质可将其划分为两个阶段,即氧化相(oxidative phase)和非氧化相(non-oxidative phase)。

一、氧化相

氧化相由3步不可逆反应组成,因有2步氧化还原反应而得名。

1. 6–磷酸葡糖的脱氢

这是一步氧化还原反应,一般被认为是磷酸戊糖途径的第一步反应。催化此反应的酶是6–磷酸葡糖脱氢酶(glucose–6–phosphate dehydrogenase, G6PD),底物6–磷酸葡糖来自于糖酵解的第一步反应,氢和电子的受体是$NADP^+$。经过这一步反应,6–磷酸葡糖被氧化为6–磷酸葡糖酸内酯(6–phosphogluconolactone),同时$NADP^+$被还原成NADPH。

$^{2-}O_3POCH_2$ … OH, HO, OH, OH　α–6–磷酸葡糖
$NADP^+$ → $NADPH+H^+$　6–磷酸葡糖脱氢酶
$^{2-}O_3POCH_2$ … OH, HO, OH, =O　6–磷酸葡糖酸内酯

这一步反应是磷酸戊糖途径的限速步骤,G6PD作为其中的限速酶,NADPH既是它的产物,又是它的强竞争性抑制剂。因此,磷酸戊糖途径完全受[NADPH]/[$NADP^+$]的相对比例控制,如果细胞内的NADPH浓度高,磷酸戊糖途径就会受到抑制,反之则被激活。

2. 葡糖酸内酯的水解

这是一步水解反应。在葡糖酸内酯酶(gluconolactonase)的催化下,葡糖酸内酯环被打开,生成6–磷酸葡糖酸。在没有酶的情况下反应也能进行,可是速率太慢,而有了内酯酶可大大加快反应的速率。

$^{2-}O_3POCH_2$ … OH, HO, OH, =O　6–磷酸葡糖酸内酯
—葡糖酸内酯酶→ (H_2O → H^+)
COO^- / HCOH / HOCH / HCOH / HCOH / $CH_2OPO_3^{2-}$　6–磷酸葡糖酸

3. 6–磷酸葡糖酸的脱氢

这是一步氧化脱羧反应,反应机制与异柠檬酸脱氢反应相似(参看第二十三章“三羧酸循环”),由

6-磷酸葡糖酸脱氢酶(6-phosphogluconate dehydrogenase)催化,氢和电子的受体仍然是 $NADP^+$。反应先是脱氢产生不稳定的3-酮-6-磷酸葡糖酸(3-keto-6-phosphogluconate)中间产物和NADPH,然后发生β-脱羧反应,生成5-磷酸核酮糖。

Quiz1 三羧酸循环中的哪一步反应与此相似?

COO^- — HCOH — HOCH — HCOH — HCOH — $CH_2OPO_3^{2-}$

$NADP^+$ → NADPH + H^+ (6-磷酸葡糖酸脱氢酶)

[COO^- — HCOH — C=O — HCOH — HCOH — $CH_2OPO_3^{2-}$]

H^+ → CO_2

CH_2OH — C=O — HCOH — HCOH — $CH_2OPO_3^{2-}$

6-磷酸葡糖酸　　3-酮-6-磷酸葡糖酸　　5-磷酸核酮糖

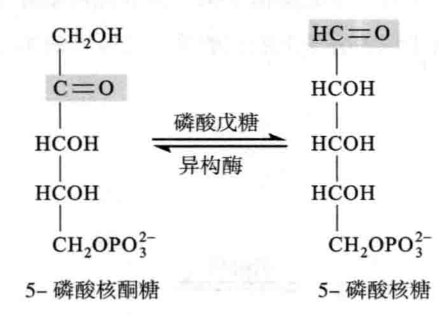
二、非氧化相

非氧化相全部由非氧化的可逆反应组成,共有5步,反应的性质是碳单位的转移、异构或分子内的重排。通过此阶段的反应,6分子戊糖转化成5分子己糖。

1. 5-磷酸核糖的形成

这是一步酮糖与醛糖进行互变的异构化反应,由磷酸戊糖异构酶(phosphopentose isomerase)催化,反应的机制涉及烯二醇中间体(参看第二十二章"糖酵解")。

Quiz2 这一步反应与糖酵解中的哪一步反应相似?



到此为止,核糖和NADPH这两种重要的物质已经形成,它们可以分别参与核苷酸的合成和其他生物分子的合成,这时的总反应式可写成:

$$6\text{-磷酸葡糖} + 2NADP^+ + H_2O \longrightarrow 5\text{-磷酸核糖} + 2NADPH + 2H^+ + CO_2$$

在某些细胞内的代谢条件下,5-磷酸核糖完全用于核苷酸的合成,几乎没有剩余,而NADPH参与其他代谢物合成,因此对于这些细胞来说,磷酸戊糖途径实际上已经结束。然而,很多细胞对于NADPH的需求远远高于5-磷酸核糖,因此这些细胞内的磷酸戊糖途径特别活跃。但与此同时,5-磷酸核糖必然出现过剩,细胞就必须存在一种机制能够及时处理这些剩余的5-磷酸核糖。事实上,磷酸戊糖途径余下的反应就是为此预备的。

Quiz3 何种细胞具有这样的特征?

余下的反应主要是分子之间碳单位的转移和分子内的重排反应,多余的戊糖将经历丙糖、丁糖、己糖和庚糖,最后变成糖酵解的中间产物进行代谢。在进行转移和重排反应之前,还需要5-磷酸木酮糖的存在,因此下一步反应就是生成5-磷酸木酮糖的反应。

2. 5-磷酸木酮糖的形成

这是一步由磷酸戊糖差向异构酶(phosphopentose epimerase)催化的异构化反应,经过这一步反应,5-磷酸核酮糖变成了它的差向异构体——5-磷酸木酮糖。

3. 第一次碳单位的转移和重排反应

这是一步由转酮酶(transketolase)催化的反应,需要TPP和Mg^{2+},反应机制类似于丙酮酸脱氢酶所催化的反应[图24-1(1)]。经过此反应,5-磷酸木酮糖分子上的二碳单位被转移到5-磷酸核糖分子上,形成3-磷酸甘油醛和7-磷酸庚酮糖(sedoheptulose)。

5- 磷酸核酮糖 $\rightleftharpoons$（磷酸戊糖差向异构酶）5- 磷酸木酮糖

5- 磷酸木酮糖 + 5- 磷酸核糖 $\rightleftharpoons$（转酮酶，TPP/Mg^{2+}）3- 磷酸甘油醛 + 7- 磷酸庚酮糖

4. 第二次碳单位的转移和重排反应

这是一步由转醛酶（transaldolase）催化的反应[图 24-1(2)]，不需要任何辅助因子，上一步反应的产物正好是这一步反应的底物，反应的机制类似于第一类醛缩酶催化的反应（参看第二十二章“糖酵解”）。经过此酶的催化，7- 磷酸庚酮糖分子上的三碳单位被转移到 3- 磷酸甘油醛分子上，形成 4- 磷酸赤藓糖和 6- 磷酸果糖。

Quiz4 如何设计一个实验证明转醛酶的催化机制类似于第一类醛缩酶的催化机制？

7- 磷酸庚酮糖 + 3- 磷酸甘油醛 $\rightleftharpoons$（转醛酶）4- 磷酸赤藓糖 + 6- 磷酸果糖

5. 第三次碳单位的转移和重排反应

这又是一步由转酮酶催化的反应，这一次 5- 磷酸木酮糖分子上的二碳单位被转移到 4- 磷酸赤藓糖分子上，形成 3- 磷酸甘油醛和 6- 磷酸果糖。

5- 磷酸木酮糖 + 4- 磷酸赤藓糖 $\rightleftharpoons$（转酮酶，TPP/Mg^{2+}）3- 磷酸甘油醛 + 6- 磷酸果糖

三、磷酸戊糖途径小结

以上介绍了一个完整的磷酸戊糖途径所涉及的所有反应，其中有以下几点需要特别注意：

（1）1 个葡萄糖分子是不可能完成上述反应的，至少要有 3 个葡萄糖分子同时进入才可以完成。

(1) 转酮酶催化的反应

(2) 转醛酶催化的反应

图 24-1 转酮酶和转醛酶催化的反应机理

(2) 与糖酵解和三羧酸循环不同之处是：1 个葡萄糖分子依次经过糖酵解和三羧酸循环后，6 个碳原子可以被完全氧化成 CO_2，但 1 个葡萄糖分子进入磷酸戊糖途径以后，只有 1 个碳被氧化为 CO_2，其余的碳原子仍然以糖的形式存在，很显然只有当 6 个葡萄糖分子同时进入磷酸戊糖途径，到最后才相当于有 1 个葡萄糖分子完全被氧化成 CO_2 和 H_2O。

(3) 磷酸戊糖途径并不是细胞产生 NADPH 的唯一途径，比如异柠檬酸脱氢酶、谷氨酸脱氢酶、苹果酸酶催化的反应和光合作用的光反应也能产生 NADPH。

(4) 磷酸戊糖途径发生在细胞质基质，不需要氧气。

(5) 与糖酵解、糖异生和三羧酸循环相比，磷酸戊糖途径的调节机制相当简单。

(6) 根据细胞对 NADPH、核糖和 ATP 的需要不同，磷酸戊糖途径可以以 4 种不同的模式存在。

Quiz5 细胞合成核糖是不是一定要有 6- 磷酸葡糖脱氢酶和 6- 磷酸葡糖酸脱氢酶？为什么？

快速分裂的细胞需要更多的 5- 磷酸核糖，以第一种模式存在；需要等量的 5- 磷酸核糖和 NADPH 的细胞以第二种模式存在；需要更多的 NADPH 以进行生物合成的细胞（如脂肪细胞）以第三种模式存在；只需要 NADPH 和 ATP，而不需要 5- 磷酸核糖的细胞以第四种模式存在（图 24-2）。

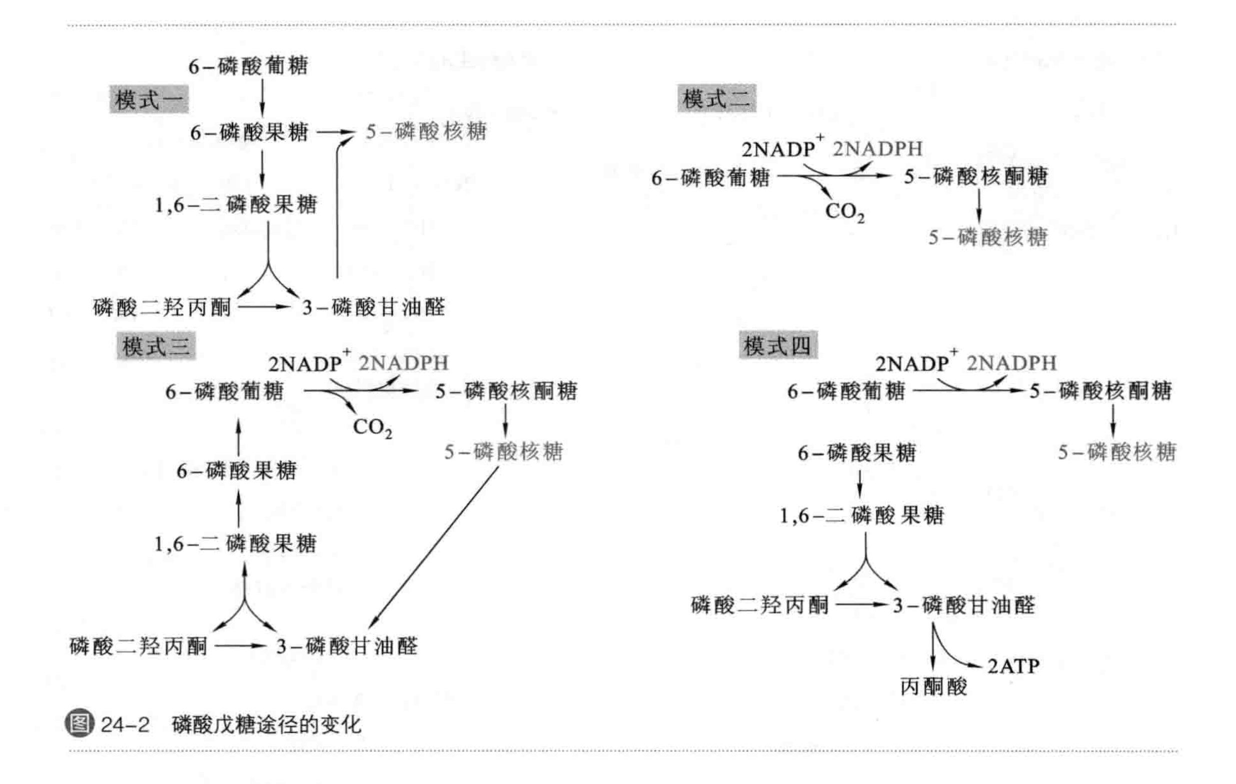


图 24-2 磷酸戊糖途径的变化

第二节 磷酸戊糖途径的生理功能

Quiz6 磷酸戊糖途径可以将戊糖转变成何种糖酵解的中间产物?

磷酸戊糖途径的功能主要与其产生的 NADPH、5-磷酸核糖和 4-磷酸赤藓糖有关。此外,它还提供了一种将戊糖转化为糖酵解中间产物的手段,从而使摄入体内的戊糖也能够最后进入糖酵解进行氧化分解。

1. 与 NADPH 有关生理功能

(1) 提供生物合成的还原剂 NADPH 需要 NADPH 提供还原能力的生物合成途径包括脂肪酸的生物合成、胆固醇的生物合成、核苷酸的生物合成和神经递质的生物合成。NADPH 的缺乏将会削弱细胞合成这些物质的能力。正因为如此,那些生物合成旺盛的组织或细胞(如肾上腺、肝、睾丸、脂肪组织、卵巢和乳腺等)内的磷酸戊糖途径就特别活跃。

e24-1 细胞色素 P450 单加氧酶解毒系统

(2) 解毒 细胞色素 P450 单加氧酶解毒系统需要 NADPH 参与对毒物的羟基化反应。

(3) 免疫 巨噬细胞(phagocyte)膜上存在一种叫 NADPH 氧化酶(NADPH oxidase)的酶,它能催化 NADPH 上的电子转移给 O_2,形成超氧阴离子以杀死入侵的微生物(图 24-3)。在巨噬细胞与微生物的"较量"中,NADPH 氧化酶需要消耗大量的 NADPH,这严重依赖于磷酸戊糖途径,因此磷酸戊糖途径的缺乏将会导致机体容易受到感染。

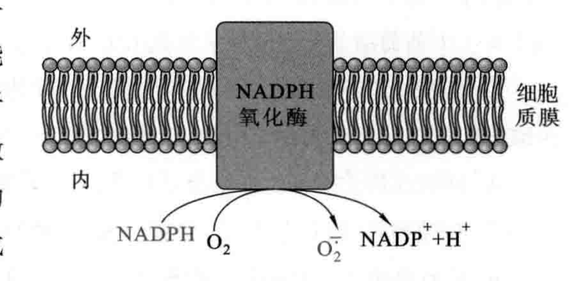


图 24-3 巨噬细胞膜上的 NADPH 氧化酶的防御功能

(4) 间接进入呼吸链 NADPH 并不能直接进入呼吸链,但在吡啶核苷酸转氢酶(pyridine nucleotide transhydrogenase)的催化下,NADPH 可将氢原子转移给 NAD^+,形成 NADH,再由 NADH 进入呼吸链产生 ATP。

(5) 维持红细胞膜的完整 NADPH 是谷胱甘肽还原酶(glutathione reductase)的辅酶,在还原型谷胱甘肽(GSH)的再生反应中发挥着重要作用(图 24-4)。

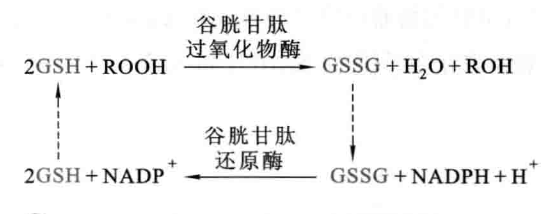


图 24-4 谷胱甘肽过氧化物酶和还原酶参与的过氧化物的清除

GSH 在红细胞中主要有两个功能：一是维护血红蛋白中的血红素铁处于二价状态；二是使红细胞膜免受过氧化物的损害。这一方面是因为红细胞富含氧气，在富氧的环境之中，极易生成对细胞膜上的脂质具有破坏性的过氧化物；另一方面是因为成熟的红细胞已经完全丧失了合成蛋白质和脂质的能力，因此如果它的细胞膜受到损害就无法修复，极易造成溶血。在 GSH 的存在下，红细胞内的一种含有 Sec 残基的谷胱甘肽过氧化物酶能及时清除过氧化物，阻止其对红细胞膜的危害，从而有利于维护膜的完整。

2. 其他有关生理功能

（1）提供核苷酸及其衍生物合成的前体 5- 磷酸核糖（参看第三十二章“核苷酸代谢”）。

（2）芳香族氨基酸和维生素 B_6 的合成需要 4- 磷酸赤藓糖。

框 24-1　生化与健康——蚕豆与溶血性贫血

蚕豆病（favism）又称为遗传性 6- 磷酸葡糖脱氢酶（G6PD）缺乏症，是一种常见的与 X 染色体连锁的不完全显性遗传病。此病患者编码 G6PD 的基因先天性缺陷，因此其体内红细胞的 G6PD 的活性很低。由于 G6PD 是 PPP 的限速酶，而 PPP 是红细胞产生 NADPH 的唯一途径，因此 G6PD 的缺陷会影响到 PPP 的活性，从而使红细胞内的 NADPH 水平极低，而无法维持还原型谷胱甘肽（GSH）的再生。

GSH 是红细胞清除引起氧化性损伤的自由基所必需的（图 24-4）。缺乏 GSH 将使红细胞在受到氧化性胁迫下很容易发生破裂。因此，缺乏 G6PD 的人在氧化胁迫状态下很容易发生溶血性贫血。氧化性胁迫源于机体感染或者机体直接接触一些富含氧化剂的药物或食物，例如抗疟疾的药物——伯氨喹（primaquine），蚕豆中的氧化剂——巢菜碱（vicine）、蚕豆嘧啶（divicine）、伴蚕豆嘧啶核苷（convicine）和异尿咪（isouramil）。这些成分在体内能够增强自由基的形成。当红细胞内的 GSH 被耗尽以后，其内的多种酶和蛋白质（包括血红蛋白）以及膜上的脂质受到氧化剂和自由基的作用而被破坏，红细胞膜的完整性因此受到影响，其中的电解质失去平衡。受损的红细胞受吞噬，离开循环系统，被隔离在脾。血红蛋白则被代谢成胆红素，高水平的胆红素可造成黄疸。

G6PD 缺乏症的临床表现与一般的溶血性贫血大致相同。因 G6PD 缺乏而诱发的重度急性溶血性贫血可使红细胞破坏过多，如不及时处理，可引起肝、肾或心功能衰竭，甚至导致死亡。有趣的是，G6PD 缺乏症患者对疟疾具有一定的抵抗能力，这是因为疟疾原虫在红细胞生存需要高水平的 NADPH。

推荐网址：

e24-2 本章小结填填看

1. http://en.wikipedia.org/wiki/Pentose_phosphate_pathway（维基百科有关磷酸戊糖途径的内容）
2. http://themedicalbiochemistrypage.org/pentose-phosphate-pathway.html（完全免费的医学生物化学课程网站，有关磷酸戊糖途径的内容）
3. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb2/part1/pentose.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关磷酸戊糖途径的内容）
4. http://www.jbc.org/content/277/50/47965.full（美国 *JBC* 杂志中一篇有关磷酸戊糖途径的综述文章）

参考文献：

1. Tozzi M G, et al. Pentose phosphates in nucleoside interconversion and catabolism. *FEBS J*, 273: 1 089–1 101, 2006.
2. Pompella A, et al. The changing faces of glutathione, a cellular protagonist. *Biochem Pharm*, 66: 1 499–1 503, 2003.
3. Foellmi L A, et al. Glucose 6–phosphate dehydrogenase: a “housekeeping” enzyme subject totissue-specific regulation by hormones, nutrients, and oxidant stress. *FASEB J*, 8: 174–181, 1994.

第二十五章 糖异生

糖异生(gluconeogenesis)泛指细胞内由乳酸或其他非糖物质净合成葡萄糖的过程。它主要发生在动物的肝(80%)和肾(20%),是动物细胞自身合成葡萄糖的唯一手段。植物和某些微生物也可以进行糖异生。

本章将重点介绍高等动物体内发生糖异生的所有反应、糖异生的生理功能,以及糖异生的调节。

第一节 糖异生所涉及的全部反应

在详细介绍糖异生的所有反应之前,首先需要明确在高等动物体内有哪些物质可以作为糖异生的原料。

可以作为高等动物糖异生代谢途径前体的物质有:乳酸、丙酮酸、奇数脂肪酸、植烷酸(叶绿素在体内代谢的中间产物)、甘油、生糖氨基酸(参看第三十一章“氨基酸代谢”)和三羧酸循环的所有中间产物。不能作为动物糖异生代谢途径前体的物质有:乙酰-CoA、偶数脂肪酸以及 Leu 和 Lys 两种严格的生酮氨基酸。

在植物和某些微生物体内,由于乙醛酸循环的存在,乙酰-CoA 或在体内能够降解产生乙酰-CoA 的物质都可以作为糖异生的前体。

下面就以丙酮酸作为起始物质,看一看它在细胞内是如何净转变成葡萄糖的。

一、从丙酮酸到葡萄糖

图 25-1 为糖异生途径和糖酵解途径的比较,从图中可以看出,糖异生并不是整个糖酵解途径简单的逆转。其原因有两个:一是糖酵解有 3 步不可逆反应;二是机体在对这两种代谢实行交互调控的时候不允许它们同时被激活或被抑制,否则就会陷入“无效”循环之中。但是,糖异生保留了糖酵解途径中的所有可逆反应(第 2 步,第 4~9 步),属于自身的反应只有 4 步。在这 4 步反应中,有 2 步是用来克服糖酵解的最后 1 步不可逆反应,其余 2 步是用来克服糖酵解的第 3 步和第 1 步

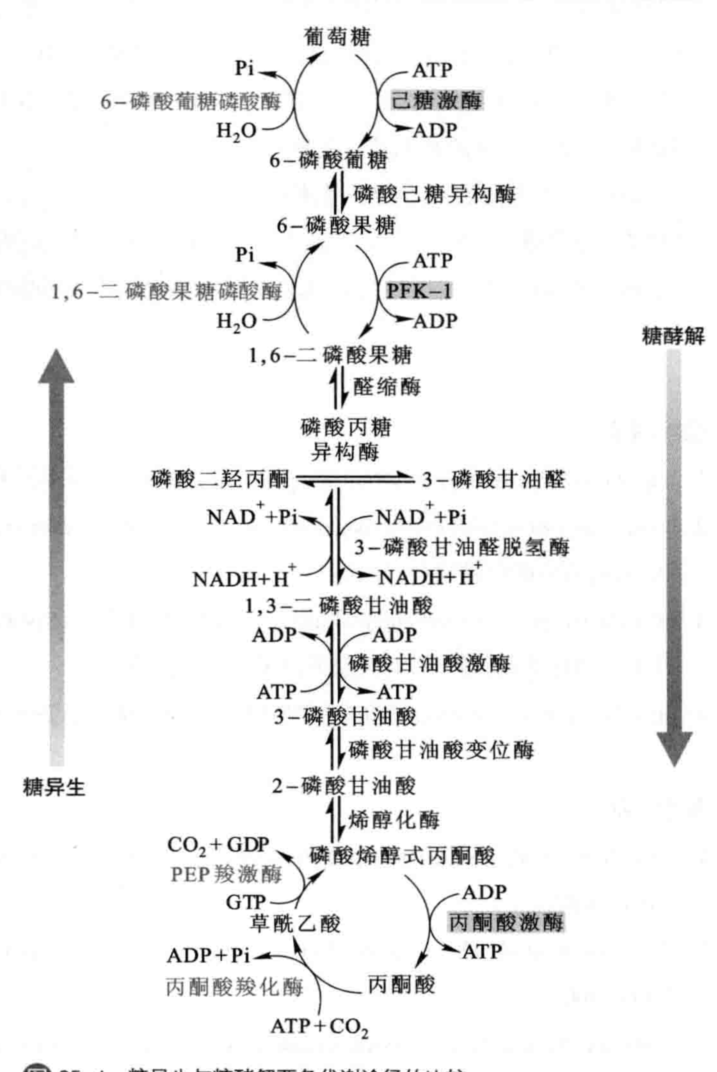


图 25-1 糖异生与糖酵解两条代谢途径的比较

不可逆反应,现分别加以讨论。

(一) 丙酮酸转变成为草酰乙酸

Quiz1 你认为碘代乙酸和氟化物能抑制糖异生吗?为什么?

这是一步发生在线粒体基质并且消耗 ATP 的羧化反应,可兼作三羧酸循环的回补反应(参看第二十三章“三羧酸循环”),由丙酮酸羧化酶催化。该酶需要生物素作为辅基,需要乙酰 -CoA 的激活。由于反应有 ATP 参与,因此还需要 Mg^{2+} 或 Mn^{2+}。

$$\text{丙酮酸} + ATP + HCO_3^- \xrightarrow[\text{生物素}]{E_1} \text{草酰乙酸} + ADP + Pi$$

$$\text{草酰乙酸} + GTP \xrightarrow{E_2} PEP + GDP + CO_2$$

丙酮酸　　草酰乙酸　　PEP

E_1:丙酮酸羧化酶;E_2:PEP 羧激酶

丙酮酸羧化酶由 4 个相同的亚基组成,每一个亚基具有 3 个结构域,它们分别具有生物素羧基载体蛋白(biotin carboxyl carrier protein,BCCP)、生物素羧化酶(biotin carboxylase)和羧基转移酶(carboxyl transferase)的活性(图 25-2)。生物素辅基以生物胞素的形式存在,它所具有的戊酸侧链与酶分子的一个 Lys 残基上的侧链共价相连,仿佛一个细长的臂(参看第十四章维生素与辅酶),从而允许它的骈环结构能在两个活性中心摆动,这对整个羧化反应至关重要。

Quiz2 三羧酸循环中有哪一种酶也具有细长的臂,以方便它在两个不同活性中心之间的摆动?

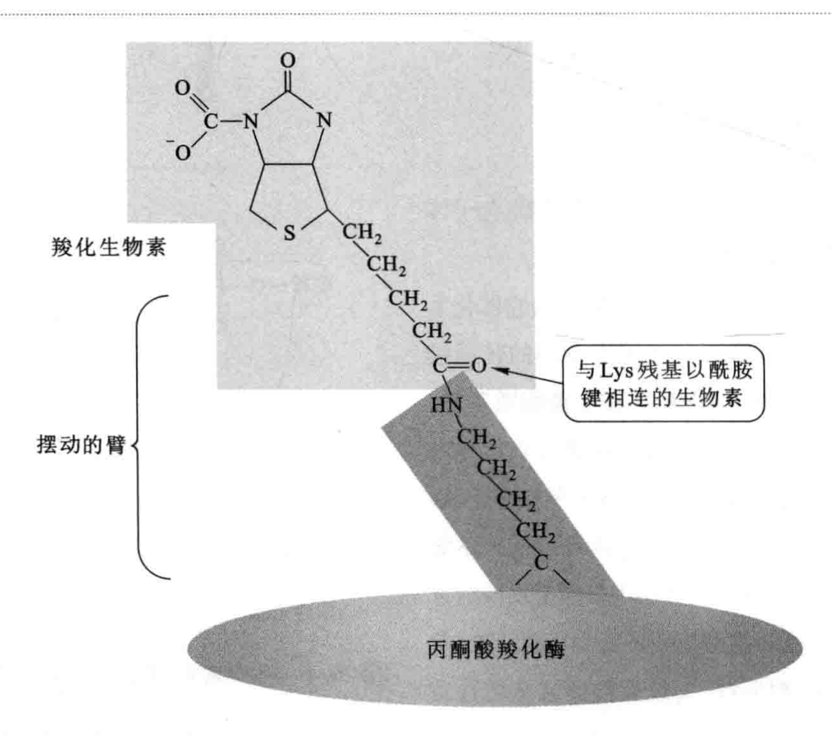


图 25-2 丙酮酸羧化酶的结构模型

整个反应分为两步(图 25-3):①生物素骈环、HCO_3^- 和 ATP 进入生物素羧化酶的活性中心,在生物素羧化酶催化下,HCO_3^- 与 ATP 形成羧基磷酸中间产物(活化的羧基),随后生物素骈环上的 N 原子接受羧基,生成羧基生物素;②羧化的生物素骈环进入羧基转移酶的活性中心,在该酶活性的催化下,羧基从生物素转移到丙酮酸分子上形成草酰乙酸。

Quiz3 如何用实验证明丙酮酸羧化酶催化的反应需要生物素?

(二) 草酰乙酸转变为磷酸烯醇式丙酮酸

这是一步消耗 GTP 的脱羧反应,丢失的羧基来源于前一步反应中刚刚被固定上去的 CO_2,由磷酸烯醇式丙酮酸羧激酶(phosphoenolpyruvate carboxykinase,PEPCK)催化,且需要 Mg^{2+} 或 Mn^{2+}。反应历程参看图 25-4,先是草酰乙酸的脱羧形成烯醇式丙酮酸阴离子中间产物,然后再发生磷酸化反应生成 PEP。

PEPCK 在不同的生物细胞内的定位不尽相同,例如在人体内,线粒体基质和细胞质基质均含有这

图 25-3 丙酮酸羧化酶的作用机理

Quiz4 有人使用基因工程的方法，在一个小鼠的骨骼肌细胞中，引入更多拷贝的 *PEPCK* 基因，使骨骼肌细胞中表达的 *PEPCK* 比正常小鼠多 10 倍，结果发现这种小鼠更加好动、侵略性更强、能跑更长的距离。对此你如何解释？

种酶，而在小鼠体内只存在于细胞质基质，兔子体内只存在于线粒体。

如果 PEPCK 存在于线粒体基质，由它催化生成的 PEP 就可以直接通过内膜上专门的转运蛋白运出线粒体；如果 PEPCK 存在于细胞质基质，就需要在线粒体基质，将不能直接透过线粒体内膜的草酰乙酸，先转变成苹果酸或天冬氨酸，然后再通过内膜上特殊的转运体，将苹果酸或天冬氨酸运出线粒体。在细胞质基质中按照逆反应的方向，将苹果酸或天冬氨酸重新转变为草酰乙酸（参看第二十二章“糖酵解”），最后再由细胞质基质中的 PEPCK 将草酰乙酸转变成 PEP。一旦 PEP 形成，就可以沿着糖酵解 6 步连续的逆反应，直到形成 1,6- 二磷酸果糖。

图 25-4 磷酸烯醇式丙酮酸羧激酶的作用机理

（三）1,6- 二磷酸果糖水解为 6- 磷酸果糖

这是一步不可逆反应，由 1,6- 二磷酸果糖磷酸酶（fructose-1,6-bisphosphatase）催化。

（四）6- 磷酸葡糖水解为葡萄糖

这也是一步不可逆反应，主要发生在肝细胞的内质网腔，由 6- 磷酸葡糖磷酸酶（glucose-6-

phosphatase）催化（参看第二十七章“糖原代谢”）。大多数细胞（如肌细胞）缺乏这种磷酸酶，因此它们的糖异生止于 6- 磷酸葡糖。

6-磷酸葡糖 $\xrightarrow[H_2O \quad Pi]{\text{6-磷酸葡萄糖磷酸酶}}$ 葡萄糖

二、其他物质进入糖异生

（一）甘油进入糖异生

甘油进入糖异生的过程类似于它进入糖酵解（参看第二十二章“糖酵解”）的过程，即先后在甘油激酶和 3- 磷酸甘油脱氢酶的催化下变成磷酸二羟丙酮，然后沿着糖酵解相反的方向，只需要克服 2 步不可逆反应就可以转变成为葡萄糖。

（二）乳酸和 Ala 进入糖异生

乳酸和 Ala 首先分别在乳酸脱氢酶和谷丙转氨酶的催化下变成丙酮酸，再由丙酮酸转变为葡萄糖。但在动物体内，乳酸发酵的场所主要是肌细胞和红细胞，而糖异生的场所主要是肝细胞。因此动物体存在一种特殊的循环，即 Cori 循环（Cori cycle）：首先在红细胞和肌细胞中产生的乳酸被释放到血液，通过血液循环运输到肝细胞，经糖异生转变为葡萄糖，然后形成的葡萄糖离开肝细胞，再经血液循环回到肌细胞和红细胞进行糖酵解氧化放能，生成乳酸（图 25-5）。

动物利用 Ala 作为糖异生的前体通常是在饥饿或禁食的情况下发生，因为在这种情况下，体内的糖原消耗很多，机体为了维持血糖浓度的稳定，使用蛋白质和脂肪作为替代能源。当肌肉细胞内蛋白质被水解为氨基酸以后，其中的 Ala 可以经过丙氨酸循环（alanine cycle）进入肝细胞（图 25-5），在肝细胞内经谷丙转氨酶的催化，转变成为丙酮酸，失去的氨基进入尿素循环，转变成尿素排出体外（参看第三十一章“氨基酸代谢”）。丙酮酸则经糖异生生成葡萄糖，葡萄糖离开肝细胞，通过血液循环回到肌细胞氧化放能，生成丙酮酸，而丙酮酸可以再转变为丙氨酸，并重复先前的循环。

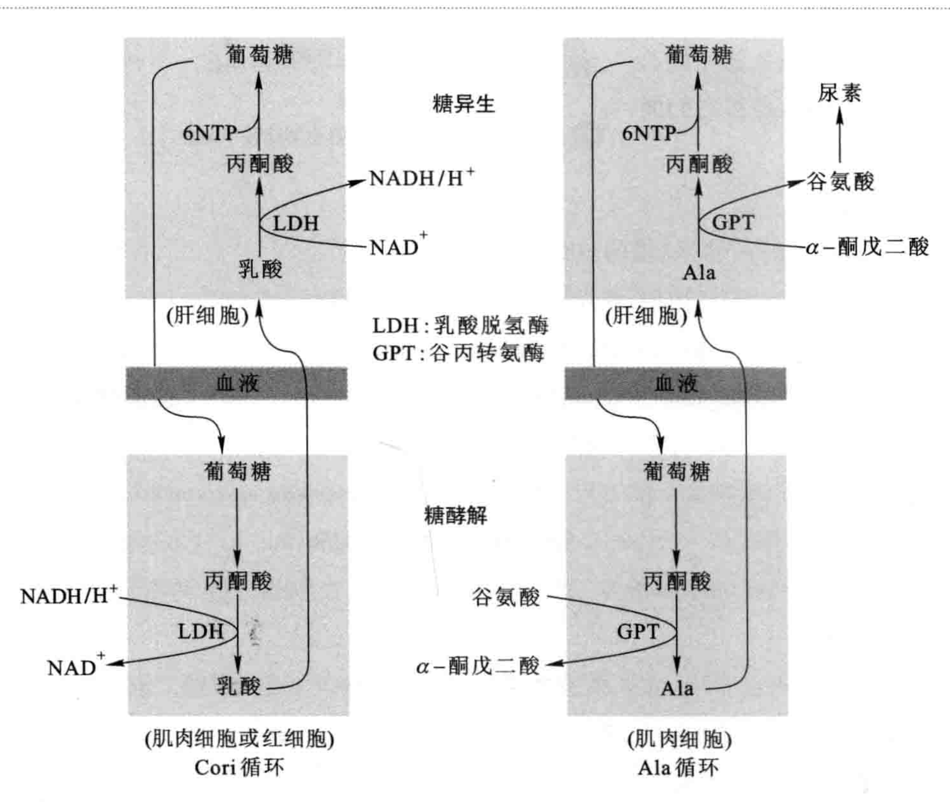


图 25-5　Cori 循环和 Ala 循环

（三）丙酸、生糖氨基酸和三羧酸循环的中间产物进入糖异生

这些物质都是通过特定的反应先转变为草酰乙酸，再进入糖异生的，具体反应分别参看第二十九章“脂肪酸代谢”、第三十一章“氨基酸代谢”和第二十三章“三羧酸循环”。

三、糖异生的能量消耗

糖异生和糖酵解的总反应的 ΔG 都是负值，因此在热力学上都是有利的。

如果以丙酮酸作为糖异生的起始物质，则总反应式为：

Quiz5 若是以草酰乙酸作为起始物质进行糖异生，每产生 1 分子葡萄糖，需要消耗多少 ATP？

$$2\text{丙酮酸} + 2NADH + 4ATP + 2GTP + 6H_2O + 4H^+ \xrightarrow{\text{糖异生}} \text{葡萄糖} + 2NAD^+ + 4ADP + 2GDP + 6Pi$$

这是一个高度耗能的过程，共消耗了 6 分子 ATP，被消耗的能量主要用来克服糖酵解途径中的 3 步不可逆反应。

第二节 糖异生的生理功能

糖异生的生理功能主要包括以下几个方面：

(1) 在饥饿或糖类摄入不足的情况下，可以补充血糖，维持血糖浓度的稳定，为那些特别依赖葡萄糖氧化放能的细胞和组织（如脑和红细胞等）提供燃料。由于糖原是一种短期能源贮备，一个人要是禁食 24 h，那么其体内所有的糖原都会耗尽，因此在长时间饥饿的情况下，糖原对稳定血糖浓度已无贡献，这时只能依赖于糖异生。乙醇在体内能够抑制糖异生，因此空腹饮酒更容易引起醉酒。

Quiz6 为什么乙醇在体内能抑制糖异生？

(2) 减轻或消除代谢性酸中毒（metabolic acidosis）。缺氧和一些疾病（如糖尿病）会导致酸性物质（如乳酸和酮体）在体内的堆积，引起代谢性酸中毒。然而，肾细胞内发生的糖异生能增强质子从体内的排出，因而能够减轻或消除代谢性酸中毒（图 25-6）。

(3) 植物和某些微生物使用乙酸作为糖异生的前体，使得它们能以乙酸作为唯一碳源。

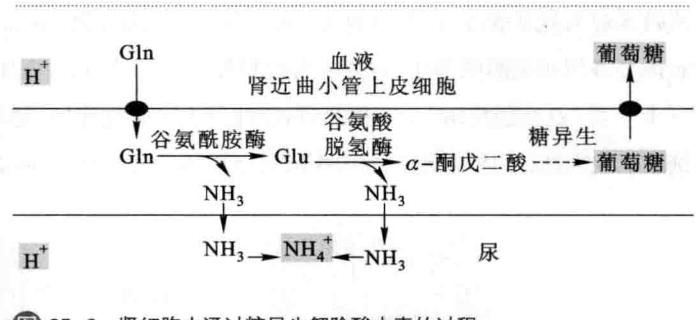


图 25-6 肾细胞内通过糖异生解除酸中毒的过程

框 25-1 生化研究动态——嗜热古菌的一种特殊的双功能酶

细菌、古菌和真核生物都能够进行糖异生，只不过是糖异生的前体会有差别。

参与糖异生的酶多数是与糖酵解共享的，如醛缩酶。然而，古菌中似乎缺乏醛缩酶。这听起来就有点不可思议！2010 年 4 月，Rafael Say 和 Georg Fuchs 发现，古菌中的 1,6- 二磷酸果糖磷酸酶在行使醛缩酶的功能，虽然参与醛缩酶反应的氨基酸残基离底物结合位点很远。

2011 年 10 月，Juan Du 和 Wei Lü 在对一种嗜热古菌（*Thermoproteus neutrophilus*）进行的研究中，获得了这种身兼两职的双功能酶——1,6- 二磷酸果糖醛缩酶 / 磷酸酶（fructose-1,6-bisphosphate aldolase/phosphatase，FBPAP）高分辨率的 3D 结构。结果他们惊奇地发现，这种酶在催化两种不同反应的时候，会重塑活性中心的结构。

FBPAP 在催化糖异生的时候，首先通过醛缩酶的活性，催化热不稳定的磷酸二羟基丙酮（DHAP）与 3- 磷酸甘油醛（GAP）缩合成 1,6- 二磷酸果糖（FBP），然后再通过 1,6- 二磷酸果糖磷酸酶的活性，将 1,6- 二磷酸果糖不可逆地水解成 6- 磷酸果糖和无机磷酸。在反应中，没有任何中间产物释放。这种酶可以通过移动环结构来重塑活性中心，使其能够催化不同的反应。

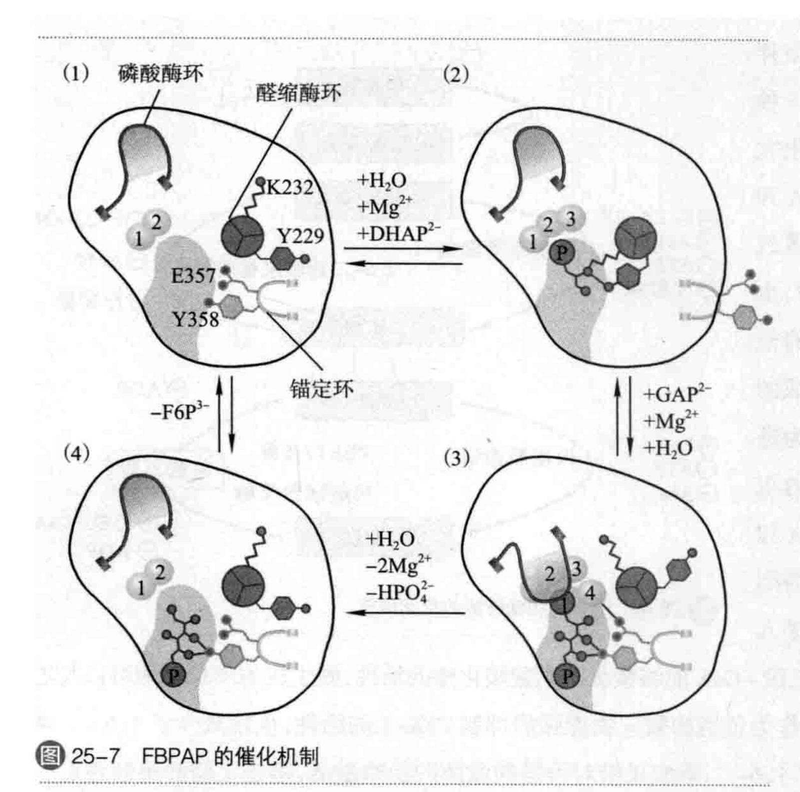


图 25-7　FBPAP 的催化机制

FBPAP 具有 3 个重要的环：第一个是醛缩酶环，由位于 220～235 位氨基酸的肽段组成，其中的 K232 和 Y229 是醛缩酶活性所必需的；第二个是磷酸酶环，由 89～111 位氨基酸的肽段组成，其作用是将中间产物 FBP 固定在结合口袋上，以方便磷酸酶的催化；第三个为锚定环，由 353～364 位氨基酸的肽段组成，其作用是将醛缩酶催化形成的 FBP 给固定住（图 25-7）。

FBPAP 在没有结合任何底物之前，酶的活性中心结合有 2 个 Mg^{2+}，底物很容易接近。当 DHAP 结合以后，第三个 Mg^{2+} 随之进入，促使醛缩酶环上的 K232 与 DHAP 的羰基形成希夫碱，这时锚定环后缩，以配合构象的变化。随后，第二个底物 GAP 进入活性中心，与 DHAP 缩合成 FBP，导致希夫碱的释放。而希夫碱的释放诱发醛缩酶环发生重排，从而创造了第四个 Mg^{2+} 的结合位点。当第四个 Mg^{2+} 结合以后，磷酸酶环发生剧烈的构象变化而将活性中心关闭。一旦 FBP 在 C1 位置的磷酸基团被水解，活性中心的四镁中心便解体，三个环立刻恢复到原来的构象和位置（图 25-7）。

嗜热古菌的两种酶融为一体，组装成活性中心可以重塑的双功能酶，显然是对高温环境的一种适应。高温下的 DHAP 和 GAP 很不稳定，FBPAP 借助构象的变化与磷酸解反应的偶联，使得醛缩酶的反应变得不可逆，这样有利于将胞内对热敏感的底物保持在较低的水平。

第三节　糖异生的调节

正如前述，糖异生和糖酵解在热力学上都是有利的。为了避免“无效”循环，生物体内的糖异生调节与糖酵解调节是高度协调的，这种协调能够保证在任何时候细胞的这两条代谢途径不可能都处于活跃的状态，而是“一张一弛”，但究竟哪一条途径被激活或被抑制完全取决于细胞当时的能量状态，以及体内葡萄糖的贮存情况和其他燃料的利用情况。

如图 25-8 所示，糖异生途径中受调节的酶，也就是用于克服糖酵解 3 步不可逆反应的 4 个酶，即丙酮酸羧化酶、PEP 羧激酶、1,6- 二磷酸果糖磷酸酶和 6- 磷酸葡糖磷酸酶。其中 1,6- 二磷酸果糖磷酸酶是最重要的调节酶，调节的方式主要是别构调节。在肝细胞，那些作为糖酵解正别构效应物的分子几乎都是糖异生的负别构效应物；相反，那些作为糖酵解负别构效应物的分子通常作为糖异生的正别构效应物。

1,6- 二磷酸果糖磷酸酶的正别构效应物是柠檬酸，负别构效应物是 2,6- 二磷酸果糖和 AMP，其中 2,6- 二磷酸果糖的浓度受到胰岛素、胰高血糖素和肾上腺素的控制（参看第二十二章“糖酵解”）。PEP 羧激酶和丙酮酸羧化酶的负别构效应物均是 ADP，丙酮酸羧化酶的正别构效应物是乙酰 -CoA。

这里需要就“机体选择乙酰 -CoA 作为丙酮酸羧化酶的激活剂”进行专门的解释：在一般情况下，

脂肪酸的氧化与糖异生是"相伴相随"的，脂肪酸作为细胞的一种高效燃料，它的氧化既可以产生大量 ATP，还可以产生乙酰 -CoA 和 NADH 等中间产物。当细胞通过脂肪酸氧化产生足够的 ATP 时，也就降低了对糖酵解产生 ATP 的依赖，这时候糖异生应该加强，生成的葡萄糖可以稳定血糖的浓度，为那些严重依赖葡萄糖氧化供能的组织提供葡萄糖。已知乙酰 -CoA 和 NADH 可以直接或间接地抑制丙酮酸的氧化脱羧，这就为丙酮酸进入糖异生提供了可能。如果此时乙酰 -CoA 能够激活丙酮酸羧化酶的活性，则生成的草酰乙酸可以与乙酰 -CoA 缩合成柠檬酸，柠檬酸作为负别构效应物能强烈抑制 PFK-1 的活性，也就减少了 1,6- 二磷酸果糖的浓度，也就等于降低了 1,6- 二磷酸果糖对丙酮酸激酶的前馈激活，阻止了糖酵解的最后一步反应。由此看来，机体选用乙酰 -CoA 作为丙酮酸羧化酶的正别构效应物，既可以协调糖异生和糖酵解，也可以协调糖代谢和脂肪酸代谢。

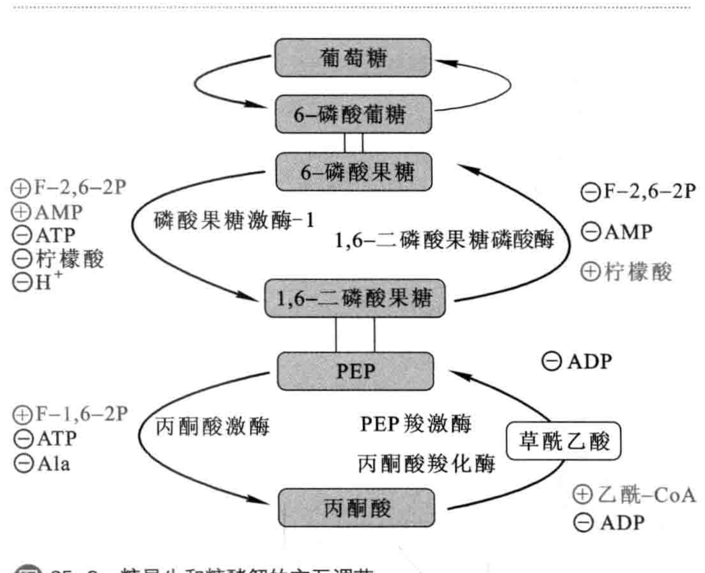


图 25-8 糖异生和糖酵解的交互调节

除了别构效应物对糖异生途径的几种调节酶的直接调控以外，还有人在研究 PEP 羧激酶的基因表达控制的时候，发现它的启动子周围具有一系列调控序列，如胰岛素应答元件（insulin-response element，IRE）、糖皮质激素应答元件（glucocorticoid hormone-response element，GRE）和 cAMP 应答元件（cAMP-response element，CRE）等（图 25-9），这说明该酶的浓度直接受到胰岛素、糖皮质激素、胰高血糖素和肾上腺素等控制。换句话说，这几种激素可以通过控制 PEP 羧激酶的基因表达直接控制糖异生。

上述所强调的糖异生和糖酵解这两种代谢途径调节的协调并不是绝对的，在某些生物体内或在某些特殊的生理条件下，细胞内参与糖酵解的限速酶和糖异生的限速酶同时被激活而发生"无效"循环。关于"无效"循环，它曾一直被视为一种浪费能量的无意义循环，但到现在为止，发现它至少具有两种功能：一是能够放大代谢调节的信号；二是能够生热。现就这两种功能分别加以讨论：如图 25-10 所示，如果 A 到 B、B 到 A 同时发生，但速率不同，其中 A 到 B 的流量为 100，而 B 到 A 的流量为 90，则 A 到 B 的净流量为 10。这时如果将 A 到 B 的流量提高 20%，B 到 A 的流量减少 20%，则原来 A 到 B 和 B 到 A 的流量分别变成了 120 和 72，净流量变成 48，这意味着速率 20% 的变化带来了 A 到 B 净流量 380% 的增加，这的确是一种可观的放大。

IRE GRE TRE CRE Ⅰ CRE Ⅱ
-500 1

图 25-9 PEP 羧激酶基因表达的调控序列

"无效"循环作为产热手段已在大黄蜂（humblebee）体内得到肯定。普通的蜜蜂很难生存在 10℃的环境下，但大黄蜂却能，其胸部

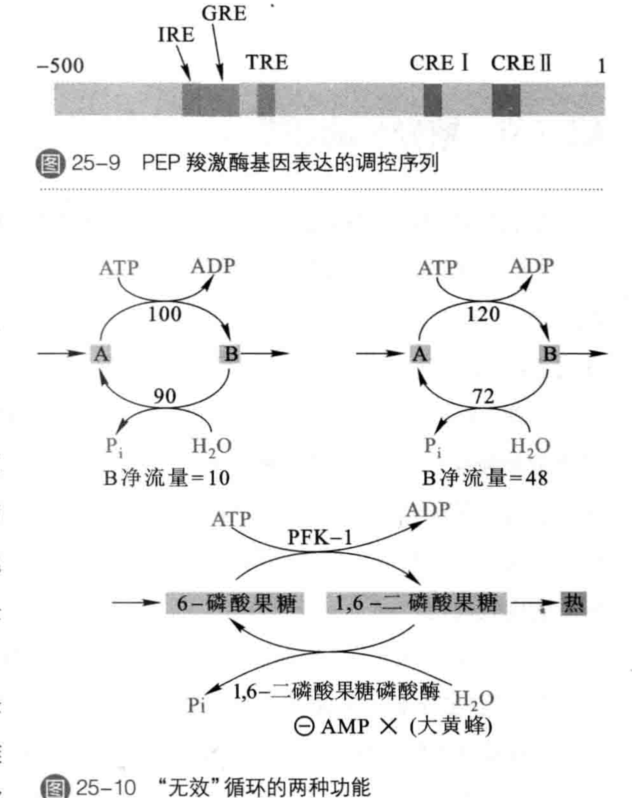


图 25-10 "无效"循环的两种功能

温度可达到30℃。而它产热的手段就是其飞翔肌细胞通过PFK-1和1,6-二磷酸果糖磷酸酶构成的“无效”循环,将ATP转变为热。已发现,它的1,6-二磷酸果糖磷酸酶不受AMP的变构抑制。显然,这种特殊的磷酸酶就是为大黄蜂产热预备的。

Quiz7 到本章为止,你知道了生物体有哪些专门产热的机制?

推荐网址:

e25-1 本章小结填填看

1. http://en.wikipedia.org/wiki/Gluconeogenesis(维基百科有关糖异生的内容)
2. http://themedicalbiochemistrypage.org/gluconeogenesis.html(完全免费的医学生物化学课程网站,有关糖异生的内容)
3. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb1/part2/gluconeo.htm(美国伦斯勒理工学院的一个分子生物化学课程网站,有关糖异生的内容)
4. http://www.wiley.com/college/boyer/0470003790/animations/gluconeogenesis/gluconeogenesis.htm(糖异生的互动动画)

参考文献:

1. Du J, et al. Active-site remodelling in the bifunctional fructose-1,6-bisphosphate aldolase/phosphatase. *Nature*, 478:534-537, 2011.
2. Montminy M, et al. A fasting inducible switch modulates gluconeogenesis via activator/coactivator exchange. *Nature*, 456:269-273, 2008.
3. Matte A, et al. Structure and mechanism of phosphoenolpyruvate carboxykinase. *J Biol Chem*, 272:8 105-8 108, 1997.
4. Hers H G, Hue L. Gluconeogenesis and related aspects of glycolysis. *Annu Rev Biochem*, 52:617-653, 1983.

第二十六章　光合作用

光合作用(photosynthesis)的字面意思就是“用光合成”,更具体的含义是指生物利用光能合成有机分子(主要是糖类)的过程。能够进行光合作用的生物统称为光合有机体(photosynthetic organism),包括绿色植物、藻类、蓝细菌(cyanobacteria)和紫色细菌。光合作用不仅是生物固定 CO_2 的主要手段,还是大气中 O_2 的主要来源。

本章将以高等植物为例,详细介绍光合作用中光能是如何被捕获、传递和转化以及 CO_2 是如何得以固定的。

第一节　光合作用的基本过程

光合作用的总反应式可写成:$CO_2 + 2H_2A \xrightarrow{光能} [CH_2O] + H_2O + 2A$

在这里可将整个反应视为一种氧化还原反应,氧化剂总是 CO_2,还原剂用 H_2A 表示,但并不总是 H_2O,也可能是 H_2S 或其他物质,这与光合有机体的性质有关(表 26–1)。如果还原剂为 H_2O,就必然有 O_2 的释放。释放出来的氧来自水分子。

Quiz1 如何设计一个简单的实验,证明释放出来的氧来自 H_2O,而不是 CO_2?

表 26–1　不同光合有机体发生的光合作用的总反应式及其还原剂

光合有机体	还原剂	总反应
植物、藻类、蓝细菌	H_2O	$CO_2 + 2H_2O \rightarrow [CH_2O] + H_2O + O_2$
绿色硫细菌	H_2S	$CO_2 + 2H_2S \rightarrow [CH_2O] + H_2O + S_2$
非硫光合细菌	H_2	$CO_2 + 2H_2 \rightarrow [CH_2O] + H_2O$
	乳酸或其他	CO_2 + 2 乳酸→$[CH_2O] + H_2O$ + 2 丙酮酸

由于绝大多数光合有机体将水作为最终的还原剂,因此总反应式常写成:

$$CO_2 + 2H_2O \xrightarrow{光能} [CH_2O] + H_2O + O_2$$

上述总反应式过于简化,并不能准确反映光合作用的过程。例如在任何光合有机体内,光并不能直接驱动 CO_2 的固定,H_2O 也不能直接还原 CO_2。整个光合作用实际上是由两个相对独立的反应,即光反应(light reaction)和暗反应(dark reaction)组成(图 26–1)。在光反应中,光能被色素捕获并用来直接驱动 H_2O 的氧化,先后有 O_2 的释放、$NADP^+$ 的还原和 ATP 的合成;在暗反应中,光反应产生的 NADPH 和 ATP 被用来驱动 CO_2 的同化。

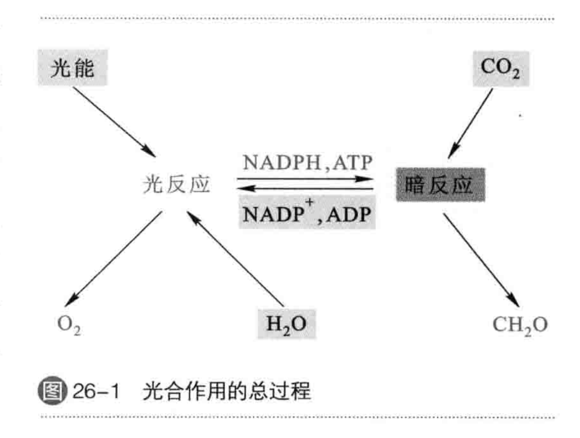


图 26–1　光合作用的总过程

第二节　植物光合作用的细胞器——叶绿体

对于高等植物和藻类植物来说,光合作用被局限在叶绿体(chloroplast)之中。叶绿体是一种由双层膜包被的细胞器,其最重要的结构特征是内部具有高度发达的片层状膜系统(图 26–2)。这可大大

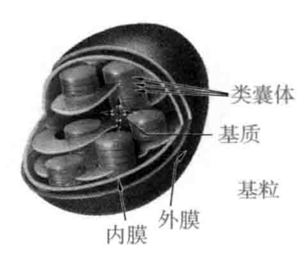
提高光吸收的面积，从而有利于光能的捕获。叶绿体主要存在于植物的叶肉细胞（mesophyll cell），其内部被内膜所包被的可溶性物质称为基质（stroma），浸入在基质中由膜包被的扁平囊状结构称为类囊体（thylakoid）。由类囊体膜包被的空间称为类囊体腔（thylakoid lumen）。由若干个类囊体垛叠在一起而构成的结构单位称为基粒（grana）。一个叶绿体内的基粒数目在 40 ~ 60，基粒之间通过基质片层（stroma lamellae）相连。在结构上，基质片层实际上是类囊体的延伸。

虽然蓝细菌和其他原核光合有机体没有叶绿体，但它们的细胞质也含有类似于类囊体的片层膜结构。

类囊体膜上集中了各种与光反应相关的成分，如光合色素、与光合色素结合的蛋白质、电子传递体、ATP 合酶等；而基质中则含有各种催化暗反应的酶、辅助因子和中间代谢产物等。因此，光合作用的光反应是在类囊体膜上进行的，而暗反应则发生在基质。

Quiz2 在高等动物体内，何种细胞具有类似的囊泡状结构？





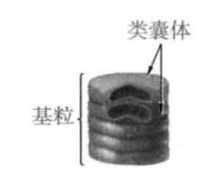


图 26-2 叶绿体的亚显微结构及其模式结构

第三节 光反应

光反应包括光能的吸收、传递和转化，水的光解，电子传递，$NADP^+$ 的还原和光合磷酸化等反应。

参与光反应的各成分在类囊体膜上组装成 5 种复合体（图 26-3），这与线粒体内膜上的呼吸链很相似。这 5 种复合体包括光系统Ⅰ（PSⅠ）、光系统Ⅱ（PSⅡ）、细胞色素 b_6/f 复合体、铁氧还蛋白 $-NADP^+$ 氧化还原酶（ferredoxin-$NADP^+$ oxidoreductase）和 CF_1F_0-ATP 合酶。前四种复合体与光能的吸收、传递、转化和电子传递以及跨类囊体膜质子梯度的建立、$NADP^+$ 的还原相关，最后一种复合体是利用质子梯度催化光合磷酸化的酶，它在结构和功能上与线粒体内膜上的 F_1F_0-ATP 合酶极为相似。在复合体之间存在流动的电子传递体，其中质体醌（plastoquinone）在结构和性质上与 CoQ 很相似，在光系统Ⅱ和细胞色素 b_6/f 复合体之间充当流动的电子传递体；质体蓝素（plasticocyanin，PC）是一种含有铜离子的可溶性膜外在蛋白，与细胞色素 c 相似，可作为细胞色素 b_6/f 复合体和光系统Ⅰ之间流动的电子载体；铁氧还蛋白则是光系统Ⅰ和铁氧还蛋白 $-NADP^+$ 氧化还原酶之间流动的电子载体。

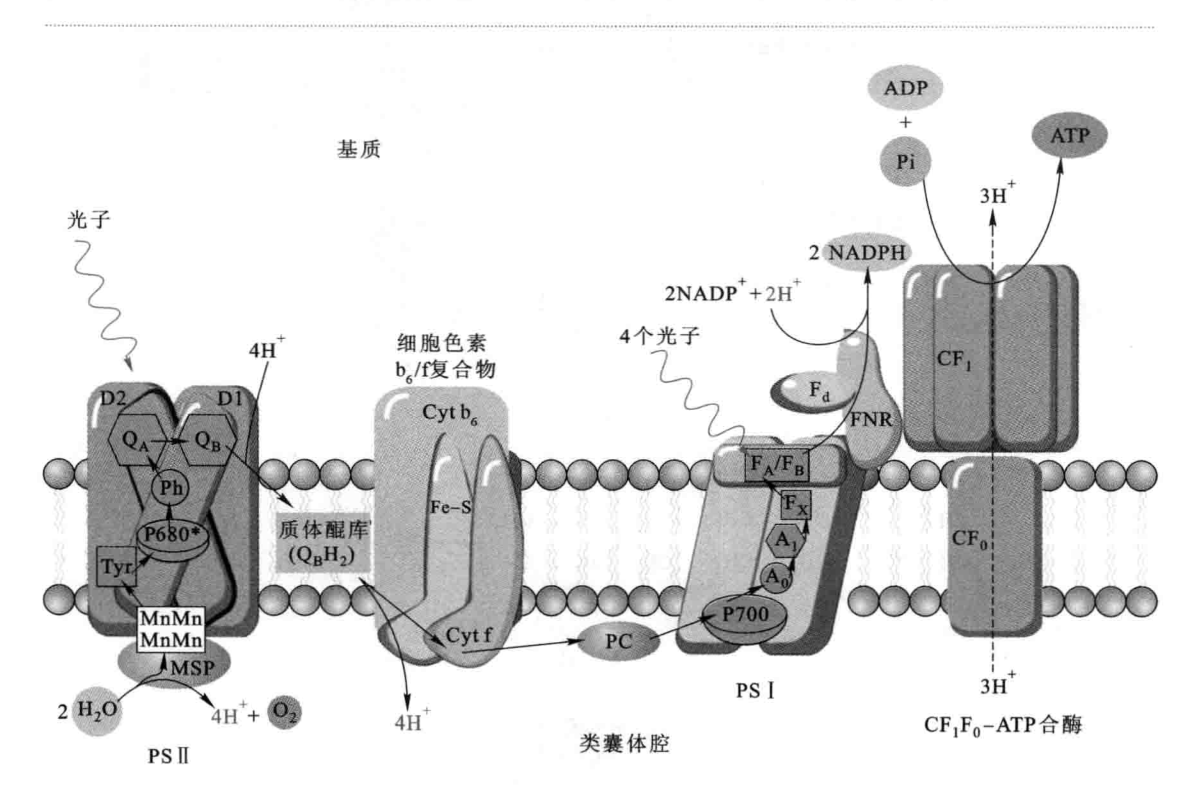


图 26-3 光反应的电子传递以及光合磷酸化

一、光能的吸收和传递

（一）光合色素的结构与功能

光合作用的第一步是光能的捕获和吸收，其最初的反应是通过类囊体膜上的吸光色素（light-absorbing pigment）来完成的。为了能够充分利用太阳光中的能量，光合有机体在进化中已形成一套由吸收波长不同的色素构成的高效的吸光装置（图 26–4）。这套装置能有效地捕获可见光和近红外光的能量。

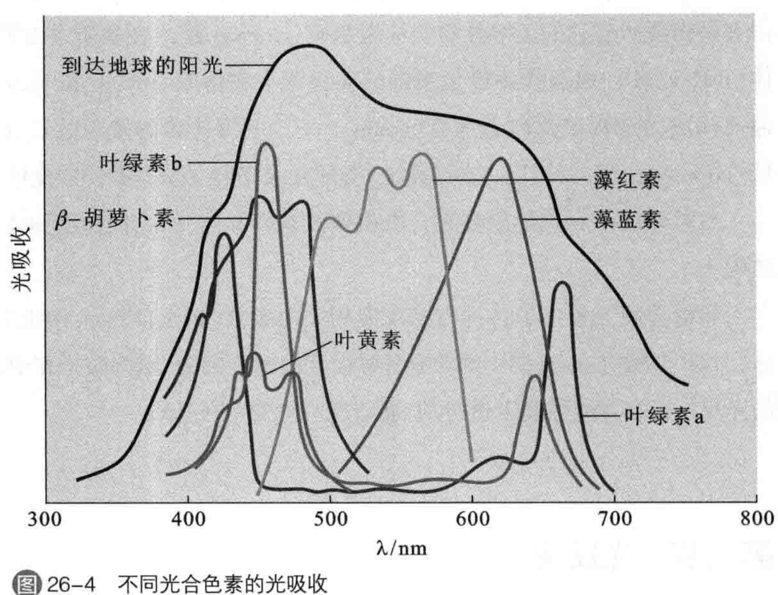


图 26–4　不同光合色素的光吸收

吸光色素主要有叶绿素（chlorophyll）、类胡萝卜素（carotenoid）和藻胆素（phycobilin）3 类，其中叶绿素又分为 a、b、c、d 和 e，类胡萝卜素包括 β- 胡萝卜素（carotene）和叶黄素（lutein），藻胆素包括藻蓝素（phycocyanobilin）和藻红素（phycoerythrobilin）。高等植物中存在的吸光色素包括叶绿素 a、b 和类胡萝卜素，藻红素只存在于蓝细菌和藻类植物，光合细菌含有细菌叶绿素（bacteriochlorophyll）。植物中的叶绿素主要吸收蓝光和红光，类胡萝卜素主要吸收蓝绿光，藻胆素主要吸收不被叶绿素和类胡萝卜素吸收的可见光，而光合细菌中的细菌叶绿素能有效地吸收红外光。绿光和黄光不能有效地被植物中的色素吸收，也正是植物叶子一般呈绿色的原因。

主要色素的化学结构见图 26–5。最值得注意的是叶绿素，它非常像血红素，但与吡咯环结合的不

叶绿素a

藻红素

图 26–5　叶绿素及藻红素的化学结构

是 Fe^{2+} 而是 Mg^{2+}。Mg^{2+} 所起的作用是影响各吡咯环上的电子分布，以保证有多个电子处在高能轨道上，从而有利于光能的捕获和转移。另外，4 个吡咯环中的 1 个已被还原。此外，叶绿素还具有一个疏水的叶绿醇链，叶绿醇链的作用是将叶绿素锚定在类囊体膜上。

虽然不同的色素在化学结构上差别很大，但它们都带有多个共轭的双键，这种特征也存在于其他一些吸光的物质之中，比如视网膜细胞中的视紫红质、紫细菌中的菌紫质和参与 DNA 紫外损伤修复的光复活酶各自的辅基都具有类似的结构。

类囊体膜上的吸光色素通常不是游离的，而是与蛋白质结合在一起，这些蛋白质不仅确保了相关色素在膜上具有合适的取向，还有助于相邻的色素之间正确地排列，减少光能在转移过程中的损失。

（二）中心色素和反应中心

色素在吸收了特定波长的光能以后，并不意味着就一定能够将吸收的光能进行转换。实际上，类囊体膜上绝大多数的色素像天线一样，仅仅起着收集光能的作用，而被统称为天线色素（antenna pigment）或辅助色素（accessory pigment）。由天线色素收集到的光能除了少数以热或荧光的形式损失以外，大多数通过共振的形式进行传递，最后被转移到一种特殊的叶绿素 a（图 26–6）。该叶绿素在类囊体膜上与特殊的蛋白质结合，被称为中心色素（central pigment）。当中心色素从天线色素获得光能以后，能量的转换反应也就开始了，这时中心色素外围处于基态（ground state）的电子发生能级跃迁，即从基态变为激发态（excitation state）。处于激发态的电子很容易失去，并转移给相邻的初级电子受体（primary electron acceptor），电荷分离（charge separation）随之发生。这种由中心色素和与它结合的蛋白质所构成的复合体被称为反应中心（reaction center）（图 26–7），而由天线色素和与之相结合的蛋白质共同组成的复合物被称为聚光复合体（light-harvesting complex，LHC）。由聚光复合体、反应中心和初级电子受体构成的光能吸收、转移和转换的功能单位称为光系统（photosystem，PS）。

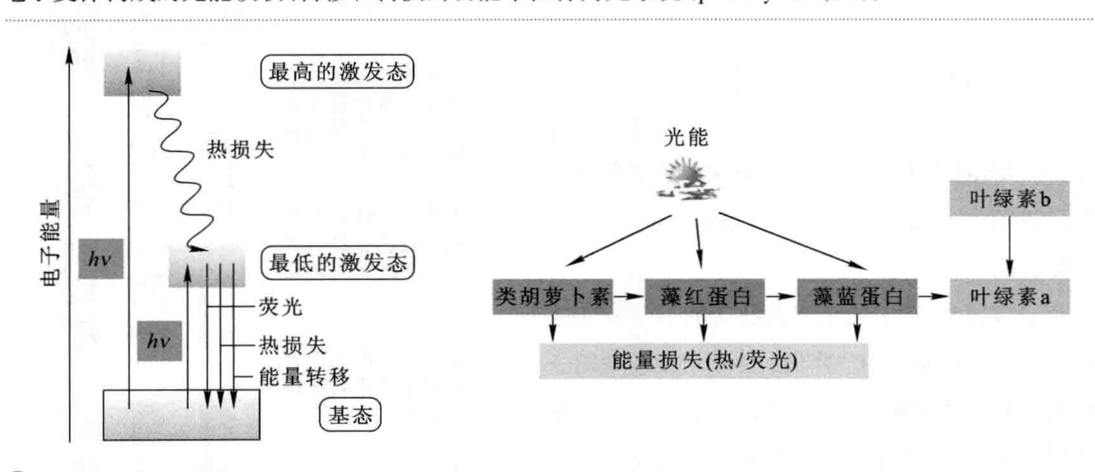


图 26–6　光能被吸收后的几种不同去处

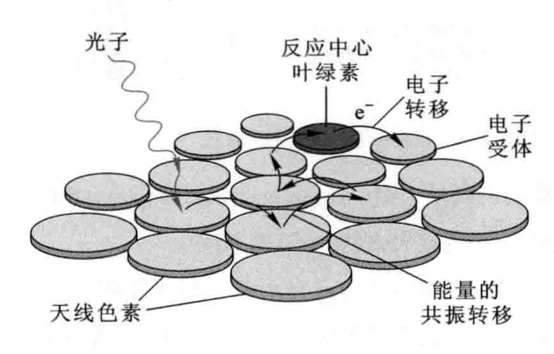


图 26–7　光合作用的反应中心

二、产氧光合有机体的光系统

在光系统的组成上，产氧光合有机体与非产氧光合有机体有显著的差别，前者具有两种光系统，即 PS Ⅰ 和 PS Ⅱ，而后者只有一种光系统。不过，各光系统在结构与功能上具有高度的同源性。

PS Ⅰ 和 PS Ⅱ 最重要的差别在于各自中心色素光吸收的最大值，前者为 700 nm（红光），因而简写为 P700；后者为 680 nm（橘红），简写为 P680。两种光系统的其他差别见表 26–2。注意该表提供的是 2011 年底科学家对嗜热细长聚球蓝细菌（*Thermosynechococcus vulcanus*）研究出来的最新数据。

表 26-2 嗜热细长聚球蓝细菌两大光系统的结构比较

光系统	PS Ⅰ	PS Ⅱ
概况	以三聚体的形式存在，总大小约为 1.02×10^6	以二聚体的形式存在，总大小约为 3.5×10^5
色素分子	100 个叶绿素 a、20 个类胡萝卜素	35 个叶绿素、2 个脱镁叶绿素、11 个胡萝卜素
中心色素	P700，每一个单体由一对特殊的叶绿素 a 组成	P680，每一个单体由一对特殊的叶绿素 a 组成
蛋白质亚基	12，如 PsaA、PsaB、PsaC、PsaD	20，如 D1、D2、CP29、CP26、CP43、CP47
金属离子	12 个 Fe^{2+} 在 3 个 Fe_4S_4 中心中，一个 Ca^{2+}	1 个非血红素 Fe^{2+}、2 个非血红素 Fe^{2+}、4 个 Mn^{2+}、3~4 个 Ca^{2+}，其中的 1 个 Ca^{2+} 与 4 个 Mn^{2+} 由 5 个 O 作为桥连成 Mn_4CaO_5 簇
其他	2 个维生素 K_1	3 个 Cl^{-1}，2 个质体醌
初级电子受体	叶绿素 a	脱镁叶绿素
电子供体	质体蓝素	水（整个二聚体有 2 795 个水分子）
功能	$NADP^+$ 的还原	水的光解、产生跨类囊体膜的质子梯度

P700 和 P680 实际上都是叶绿素 a，同一种色素光吸收却不一样的原因是它们分别与不同的蛋白质结合在一起，即所处的微环境是不同的。

多数光系统内天线色素的数目并不是固定不变的，生长在阴暗环境下的光合有机体一般比生长在光线较强的环境中的光合有机体多。

图 26-8 显示了一个光系统捕捉到光能以后所发生的反应，反应从天线色素捕获光子开始，以电荷分离结束。

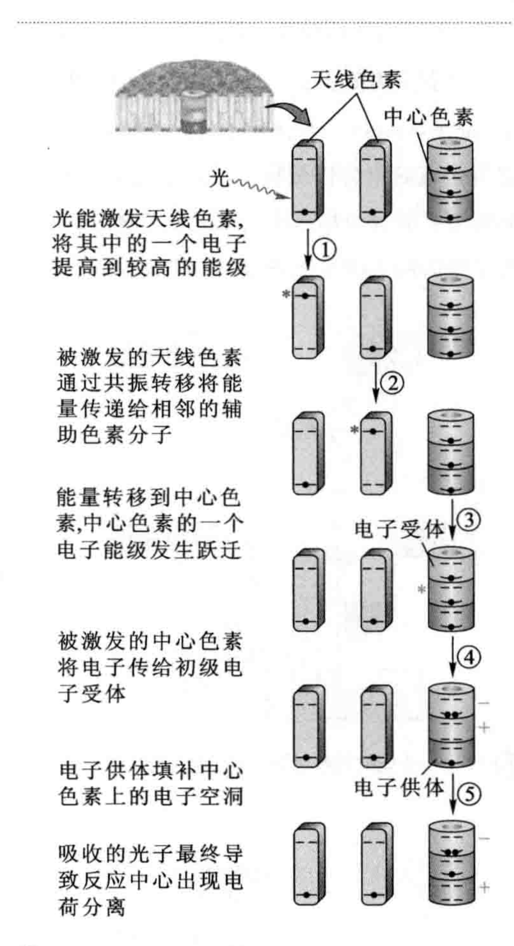


图 26-8 反应中心内发生的光反应

三、PS Ⅱ内发生的光化学反应

光反应开始于 PS Ⅱ，其一对中心色素（P680）与几种电子传递体和 2 个分别叫 D_1、D_2 的蛋白质结合在一起，所催化的总反应式为：$2H_2O + 2PQ_B + 4$ 光子 $\longrightarrow O_2 + 2PQ_BH_2$

分步叙述为：①天线色素捕获光能→共振传递到反应中心→ P680 的 $E^{\ominus}$降低→ 1 个电子被激发→转移到与 D_1 结合的脱镁叶绿素（pheophytin，Ph）。脱镁叶绿素属于一种被修饰的叶绿素 a，它与正常叶绿素的差别在于与吡咯环结合的 1 个 Mg^{2+} 被换成了 2 个质子。②被还原的 Ph 将得到的 1 个电子传给与 D_2 结合的质体醌（Q_A）。③先后有 2 个电子从 Q_A 传给与 D_1 结合的质体醌（Q_B），Q_B 同时从基质夺取 2 个质子，被还原成 Q_BH_2。④由 4 个 Mn^{2+}、若干 Ca^{2+} 和若干锰稳定蛋白（manganese stabilizing protein，MSP）构成的产氧复合物（oxygen-evolving complex，OEC）（也称为锰中心，Mn center）催化水的光解（photolysis），补充中心色素失去的电子（图 26-9）。

具体反应是：$2H_2O$ → 4 个电子→ D_1 上的 Tyr 残基接受 1 个电子→ Tyr 自由基→氧化型的 P680 得到电子被还原。与 MSP 结合的 4 个 Mn^{2+} 能够积累 4 个电子，就像呼吸链上的复合体Ⅳ，可有效地防止一些有毒的部分氧化产物（如超氧阴离子）的形成。

水的光解是在类囊体腔内发生的，在这一步有 4 个 H^+ 和 1 分子 O_2 被释放到类囊体腔。

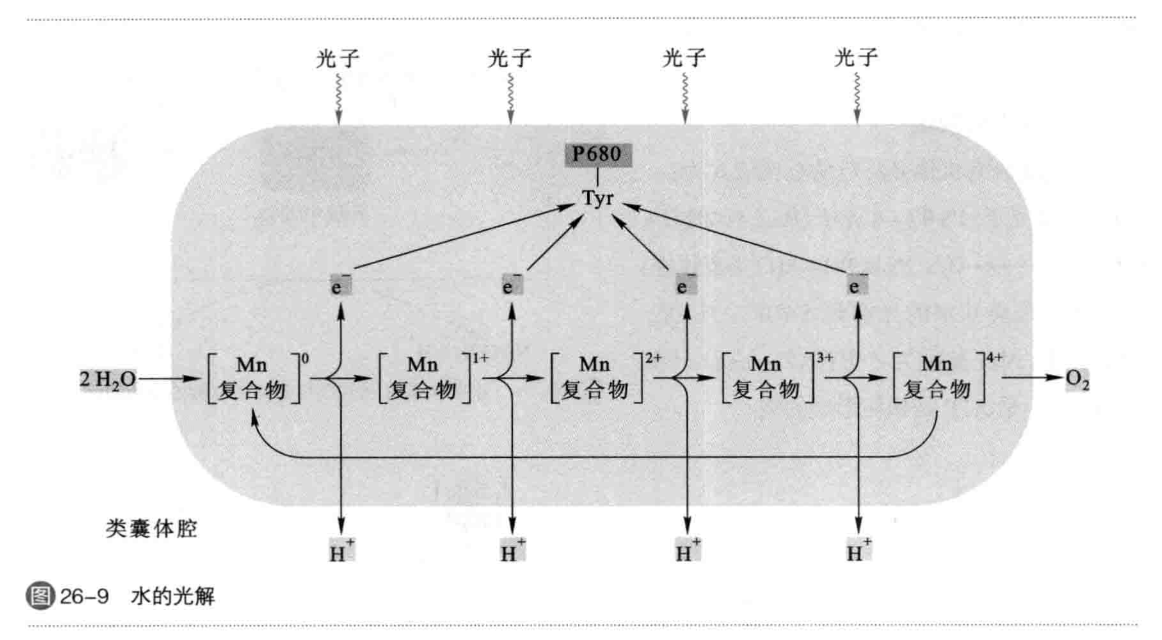


图 26-9　水的光解

四、电子从 PS Ⅱ经细胞色素 b_6/f 复合体到 PS Ⅰ的传递

在结构和功能上，细胞色素 b_6/f 复合体与呼吸链上的复合体Ⅲ十分相似。发生的总反应式是：

$$PQ_BH_2 + 2PC(Cu^{2+}) \longrightarrow PQ_B + 2PC(Cu^{+}) + 2H^{+}(类囊体腔)$$

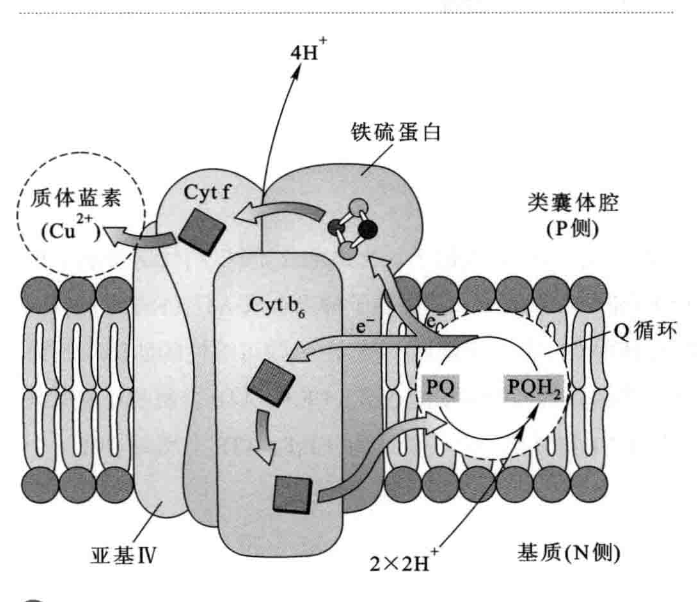


图 26-10　b_6/f 复合体中的电子传递

详细步骤为(图 26-10)：①PQ_BH_2 与 D_1 解离并流动到细胞色素 b_6/f 复合体；②PQ_BH_2 前后分两次将 2 个电子依次通过细胞色素 b_6、铁硫蛋白和细胞色素 f 传到流动的电子载体——质体蓝素(PC)，PC 上的 Cu^{2+} 被还原成 Cu^{+}；③传递的过程类似于呼吸链上的 Q 循环，有 2 个质子被释放到类囊体腔，作为对质子梯度贡献的一部分。④被还原的 PC(Cu^{+})离开细胞色素 b_6/f 复合体，进入 PS Ⅰ，补充 P700 失去的电子。

五、PS Ⅰ内发生的光化学反应

在 PS Ⅰ 内，一对中心色素(P700)与几种电子传递体、一对结构相似的蛋白质(psaA 和 psaB)紧密结合，所催化的总反应式为：

$$2PC(Cu^{+}) + 2Fd(Fe^{3+}) \longrightarrow 2PC(Cu^{2+}) + 2Fd(Fe^{2+})$$

详细步骤如下：①天线色素捕获光能→共振传递到反应中心→ P700 的 $E^{\ominus\prime}$ 降低→ 1 个电子被激发→转移到 A_0(1 个叶绿素分子)→ A_1(phylloquinone，叶绿醌)→ 3 个铁硫蛋白(FX、FA 和 FB)。②FA/FB 将电子传给铁氧还蛋白。③氧化型的 P700 从还原型的 PC 补充电子。

Quiz3　如果让两个光系统只接受红光照射，是不是两个光系统都可以被激发？为什么？

六、$NADP^+$ 的还原

反应由铁氧还蛋白 -$NADP^+$ 氧化还原酶催化，总反应式为：

$$4Fd(Fe^{2+}) + 2NADP^{+} + 2H^{+}(基质) \longrightarrow 2NADPH + 4Fd(Fe^{3+})$$

铁氧还蛋白 -$NADP^+$ 氧化还原酶是一种黄素蛋白，FAD 是它的辅基，反应机理见图 26-11。

由于反应发生在类囊体膜的基质一侧(N 侧)，因此基质内少了 2 个质子，也就相当于类囊体腔(P 侧)

多了 2 个质子，这是对质子梯度的又一贡献。

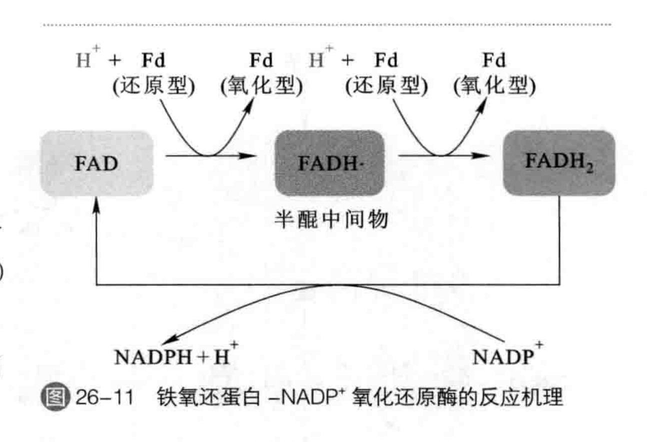


图 26-11 铁氧还蛋白 $-NADP^+$ 氧化还原酶的反应机理

七、光合磷酸化

产氧光合有机体光反应的总反应式为：$NADP^+$ + 4 光子(PS Ⅱ) + 4 光子(PS Ⅰ) + $2H_2O$ + $6H^+$(基质) ⟶ O_2 + 2NADPH + $8H^+$(类囊体腔)

整个反应从水的光解到 $NADP^+$ 的还原，形如 Z 字，因此被称为 Z 形图(Z scheme)(图 26-12)。但是 ATP 是如何形成的呢？

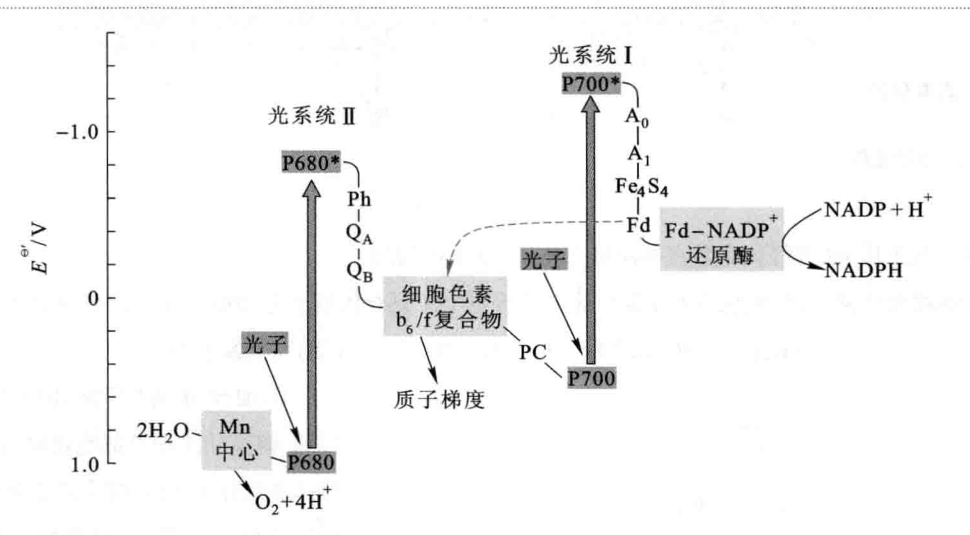


图 26-12 光反应的 Z 形图

在光反应中合成 ATP 的方式称为光合磷酸化，其机理类似于氧化磷酸化，可以用化学渗透学说进行解释，同样对解偶联剂(如 DNP 和 FCCP)和寡霉素敏感。利用质子梯度合成 ATP 的酶为 CF_1F_0-ATP 合酶，明显地分为 CF_1 和 CF_0 两个部分，其中 CF_1 突出于基质一侧，由一些可溶性的肽链组装而成($\alpha_3\beta_3\varepsilon\delta\gamma$)；而 CF_0 横跨类囊体膜，为质子从类囊体腔回到基质的通道。CF_1F_0-ATP 合酶的作用机制也可以用结合变化学说来解释(参看第二十章“生物氧化”)。一般认为，CF_1F_0-ATP 合酶每合成 1 个 ATP 需要消耗 4 个质子。

八、光反应的调节

光反应中最重要的事件之一是电子在Z形图上的流动，这种流动是受到调节的，调节有两种方式：一种是通过对 LHC 的“可逆磷酸化”改变 LHC Ⅱ与 PS Ⅱ的结合情况，加快电子从 PS Ⅱ流向 PS Ⅰ的速率；另外一种方式是抑制 PS Ⅱ或者干脆关闭 PS Ⅱ，让电子通过循环式磷酸化产生更多的 ATP。

(一) LHC Ⅱ的“可逆磷酸化”

PS Ⅰ和 PS Ⅱ在空间上是相对独立的，它们分布在类囊体膜上不同的区域。PS Ⅱ主要集中在垛叠的膜上，而 PS Ⅰ和 CF_1F_0-ATP 合酶主要分布在非垛叠的基质片层膜上。这种特殊的安排能够保证 LHC Ⅱ主要结合并激活 PS Ⅱ，而与 PS Ⅰ保持一定的距离，从而防止 LHC Ⅱ收集到的光能进入 PS Ⅰ而被“盗用”。因为如果 PS Ⅰ和 PS Ⅱ相邻，LHC Ⅱ收集到的高能量的短波长光就既可以激活 PS Ⅱ中的 P680，也可以激活 PS Ⅰ中的 P700。但 LHC Ⅰ收集到的低能量的长波长光(相对于 LHC Ⅱ)却不能激活 PS Ⅱ中的 P680(图 26-13)。

但是，LHC Ⅱ与 PS Ⅱ的结合并非固定不变，而是可以调节的，调节的方式为蛋白质的“可逆磷酸化”。如图 26-14 所示，在强光照射下，PS Ⅱ比 PS Ⅰ吸收的光能高，这导致 PS Ⅱ更为活跃，PS Ⅰ难于

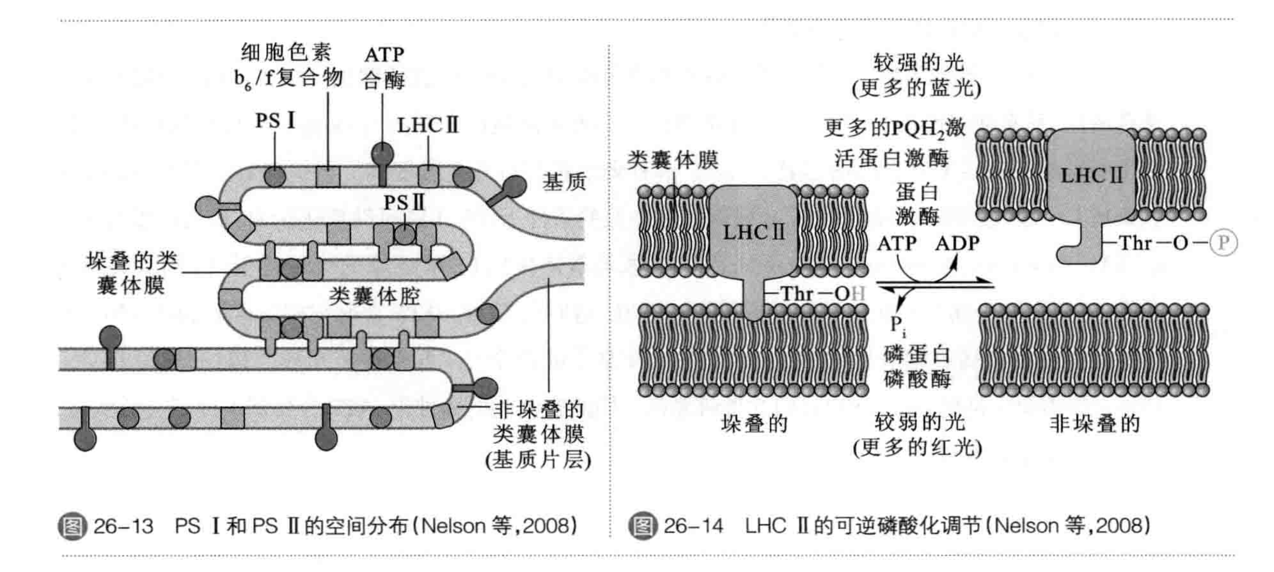


图 26-13 PS Ⅰ和 PS Ⅱ的空间分布(Nelson 等,2008)

图 26-14 LHC Ⅱ的可逆磷酸化调节(Nelson 等,2008)

跟上 PS Ⅱ的节奏,PQH_2 出现堆积转而激活一种特殊的蛋白质激酶,LHC Ⅱ上的 1 个 Thr 残基因此被磷酸化修饰。一旦 LHC Ⅱ发生磷酸化,即与 PS Ⅱ解离,转而为 PS Ⅰ提供光能,PS Ⅰ因此而加快,并与 PS Ⅱ同步。当日光减弱,堆积的 PQH_2 逐渐消失,叶绿体内的磷蛋白磷酸酶被激活,LHC Ⅱ上的磷酸基团被水解,LHC Ⅱ重新与 PS Ⅱ结合,专为 PS Ⅱ提供光能。

(二) 循环式光合磷酸化

光反应中形成的 ATP 和 NADPH 之间的比例并不是恒定的,机体可以根据需要对两者的产量实时进行调整。调整的方式是通过循环式光合磷酸化(cyclic photophosphorylation)来增加 ATP 的产量。在循环式光合磷酸化中,被还原的 Fd 并没有将电子交给 $NADP^+$,而是交给细胞色素 b_6/f 循环使用,以产生更大的质子梯度。某些光合细菌在缺乏葡萄糖的情况下,就干脆关闭 PS Ⅱ,以生成更多的 ATP。

框 26-1 生化新发现——分子变光开关

绿色植物在光合作用的时候,能量转移的效率可达 97%。在这个能源日益短缺的时代,科学家一直在梦想着有一天能够建立一种与天然光合系统相媲美的人工光合系统,从而为人类获得终极的绿色可再生能源做出贡献。然而,任何人工的聚光系统都可能面临一个问题,就是在光能过载的时候会发生某种形式的损坏。

事实上,大自然中的绿色植物每天都会面临同样的问题:即在光合作用期间,若是吸收过多光能,如何能够保护自己,避免氧化损伤呢? 早在 2005 年,以美国加利福尼亚州立大学伯克利分校的 Graham Fleming 为首的研究小组就发现一种叫玉米黄素(zeaxanthin)的类胡萝卜素,它能充当绿色植物光保护的安全出口。在两大光系统中,强光的照射可诱导 PS Ⅱ中玉米黄素的形成。形成的玉米黄素与激活的叶绿素分子相互作用,通过电荷转移(charge-transfer)机制,消耗掉多余的能量。这种机制就是玉米黄素将电子丢给叶绿素而形成自由基阳离子,而叶绿素的能量状态因此回到基态。与被激活的叶绿素不同,成为自由基阳离子的玉米黄素并非一种氧化剂。但叶绿素和玉米黄素之间的相互作用是如何调控的呢?

2008 年,Fleming 和他的研究小组在绿色植物体内又发现了一种分子变光开关(dimmer switch),能够通过聚光蛋白系统控制光能的流向。在聚光蛋白中,有一种叫 CP29 的色素结合蛋白被发现充当一种阀门,控制光合作用中多余光能的释放。而控制这种分子阀门开启的是其周围 pH 的变化。这是一种能量淬灭(energy-quenching)的光保护机制,它根据 pH 的变化检测到多余的能量,并将多余的能量从一个分子转移到另一个分子,使其被安全地消耗掉。这种防护机制对光条件的变化特别敏感,即使一阵浮云飘过也会做出反应。

调节叶绿素和玉米黄素作用的是 3 种小的聚光蛋白——CP29、CP26 和 CP24。通过对 C29 的专门研究,发现过多光能引起的 pH 变化可导致 CP29 的构象发生变化,而 CP29 的构象变化又引发受激子偶联的叶绿素的氧化还原电位发生变化。这便促进了它与玉米黄素之间的电荷转移,而产生自由基阳离子。

(三) CF_1F_0-ATP 合酶活性的调节

CF_1F_0-ATP 合酶的活性主要受到两种方式的调控:①受 pmf 的直接调控。CF_1F_0-ATP 合酶并非总是有活性,只有在足够的 pmf 下,它才能被激活。②硫氧还蛋白介导的 γ 亚基上巯基的氧化和还原,而改变激活 CF_1F_0-ATP 合酶所需要的 pmf。这种调节机制对光高度敏感。与这种机制有关的是位于 γ 亚基上一段九肽序列中的两个 Cys 残基。在光照的条件下,PS Ⅰ通过铁氧还蛋白:硫氧还蛋白氧化还原酶(ferredoxin:thioredoxin oxidoreductase)还原硫氧还蛋白 f。被还原的硫氧还蛋白 f 再去还原 γ 亚基,使两个 Cys 残基上的巯基以还原的形式存在,这时的 CF_1F_0-ATP 合酶在相对低的 pmf(>50 mV)下就能被激活。而在暗适应约 10 min 以后,γ 亚基上的两个 Cys 残基被氧化成二硫键,这时 CF_1F_0-ATP 合酶需要更高的 pmf(~100 mV)才能被激活。因此在无光的条件下,ATP 合酶基本处于失活状态,这样可以防止其水解 ATP。

第四节　暗反应

暗反应是以循环的方式进行的,它与光反应的比较参看表 26-3。循环的第一步反应是 CO_2 的固定,但 CO_2 的最初受体依植物的种类不同而异,小麦和水稻等大多数植物中 CO_2 的最初受体是一种五碳糖。当 CO_2 被固定以后,经过一个不稳定的六碳中间产物后迅速裂解为三碳中间产物,因此将发生在这些植物体内的循环称为 C_3 循环(the C_3 cycle),而进行 C_3 循环的植物统称为 C_3 植物。而在玉米和甘蔗等植物体内,CO_2 的最初受体为三碳糖,被固定后形成的是四碳中间产物,因而将这种循环称为 C_4 循环(the C_4 cycle),能进行 C_4 循环的植物称为 C_4 植物。

► 表 26-3　光反应与暗反应的比较

	光反应	暗反应
反应场所	类囊体膜	基质
主要底物	H_2O、$NADP^+$、ADP	NADPH、CO_2、ATP
主要产物	O_2、NADPH、ATP	糖类
反应动力	光能	ATP、NADPH
反应时间	光照下	有光、无光均可进行,但没有"光反应","暗反应"也不能进行
影响速率的因素	光强度和光波长	温度、CO_2 和 O_2 的浓度
能量转换	光能转变为化学能	化学能之间的转变

一、C_3 循环与 C_3 植物

Quiz4 ▶ 蔗糖和淀粉合成的场所在哪里?

C_3 循环主要是由美国生物学家 Melvin Calvin 所阐明(详见本章科学故事),因此又被称为卡尔文循环(Calvin cycle)。整个循环分为固定、还原和再生 3 个阶段(图 26-15),现分述如下:

1. CO_2 的固定

这是一步由 1,5- 二磷酸核酮糖羧化酶(ribulose-1,5-biphosphate carboxylase,RUBP 羧化酶)催化的反应,1,5- 二磷酸核酮糖(ribulose-1,5-biphosphate,RUBP)为 CO_2 的受体,该羧化酶需要 Mg^{2+},但不需要生物素,反应的产物为 2 分子 3- 磷酸甘油酸,反应机制牵涉到烯醇式中间体,共分为 5 步(图 26-16):①从 C3 抽取一个质子,产生被酶稳定的烯醇式中间产物。②CO_2 固定到 C2 上,产生 β- 酮酸中间产物(β-keto acid intermediate),即 2- 羧基 -3- 酮 -1,5- 二磷酸阿拉伯糖醇(2-carboxy-3-keto-arabinitol-1,5- bisphosphate)。③水进攻 C3 产生水合的中间产物。④水合中间产物裂解成两个部分,一部分直接是磷酸甘油酸,另一部分为磷酸甘油酸碳负离子。⑤质子快速转移给磷酸甘油酸碳负离子,产生又一个磷酸甘油酸。

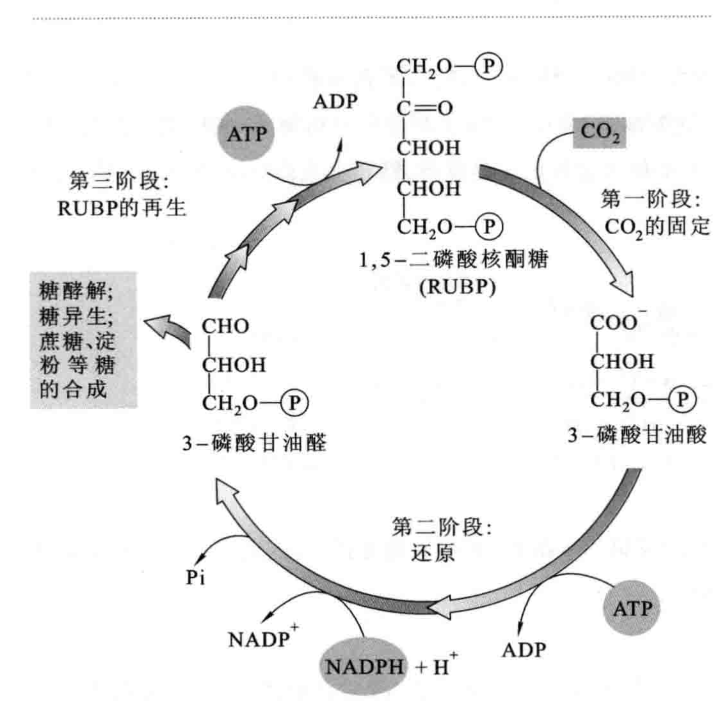


图 26-15　暗反应的全过程

由于 RUBP 羧化酶还兼有加氧酶(oxygenase)的活性,因此它经常被简写为 Rubisco。光合细菌和绿藻的 Rubisco 分别属于二聚体和八聚体蛋白。高等植物的 Rubisco 则是一种十六聚体蛋白,由 8 个大亚基(L)和 8 个小亚基(S)组成,因此可写成 L_8S_8。

Rubisco 的作用非常重要,但其催化的效率并不高,简直是低得出奇!一种典型的酶每秒钟能转化 1 000 个底物分子,但 Rubisco 每秒钟仅仅能固定约 3 个 CO_2。植物细胞通过合成大量的 Rubisco 来弥补其催化效率的先天不足。可以说,叶绿体基质内充满着 Rubisco,约占总蛋白的一半。这使得它成为地球上含量最丰富的蛋白质。

RUBP　　烯醇式中间产物　　2-羧基-3-酮-1,5-二磷酸-D-阿拉伯糖醇　　水合中间产物　　3-磷酸甘油酸

图 26-16　RUBP 羧化酶催化反应的机理

Rubisco 的另外一个"先天不足"是其特异性不高。虽然 O_2 和 CO_2 在形状和化学性质上相似,但那些与氧气结合的蛋白质(如血红蛋白)很容易排除 CO_2,因为 CO_2 比 O_2 大。而在 Rubisco 分子上,O_2 能很好地结合在为 CO_2 设计的结合位点上。因此,Rubisco 能够将 O_2 结合到糖链上,形成一种有缺陷的氧合产物,导致产生光呼吸。针对这种情况,植物细胞必须发展一套补救反应,以尽可能纠正错误。

高等植物 Rubisco 的每 2 个大亚基通过反平行的 β 折叠(由一个单体的 N 端结构域和另一个单体 C 端结构域)组装成二聚体,每 1 个二聚体在其单体接触的界面具有 2 个活性中心,因此 1 个 Rubisco 共有 8 个活性中心(图 26-17)。小亚基位于大亚基二聚体的两端,其功能可能是用来改变酶的动力学性质,使酶对温度、CO_2 浓度和 O_2 的变化做出反应。

高等植物的 Rubisco 大亚基由叶绿体基因组编码,而小亚基由细胞核基因组编码。这可以作为叶绿体内共生起源学说的一个重要证据。

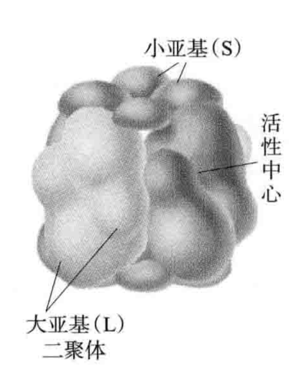
2. 3-磷酸甘油酸的还原

由 2 步反应组成,分别由 3-磷酸甘油酸激酶和 3-磷酸甘油醛脱氢酶催化,各自消耗了 ATP 和 NADPH。糖酵解也有这 2 步反应,但那里的 3-磷酸甘油醛脱氢酶使用 NAD^+ 为辅酶。

3-磷酸甘油醛作为卡尔文循环的中间产物,一部分离开循环,既可以作为合成其他糖类(葡萄糖、蔗糖、纤维素和淀粉)的原料,也可以为氨基酸和脂肪酸的合成提供碳源;另一部分则继续留在循环中用于 RUBP 的再生。

3. RUBP 的再生

参与这一个阶段反应的酶有：磷酸丙糖异构酶、醛缩酶、1,6- 二磷酸果糖磷酸酶、二磷酸景天庚酮糖磷酸酶（sedoheptulose bisphosphatase）、转酮酶、转醛酶、磷酸戊糖差向异构酶、磷酸核糖异构酶和磷酸核酮糖激酶（phosphoribulokinase）。这些酶绝大多数已在糖酵解、糖异生或磷酸戊糖途径中见过，但却是由不同的基因编码的。



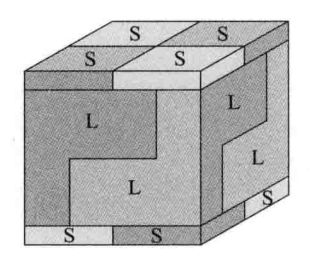
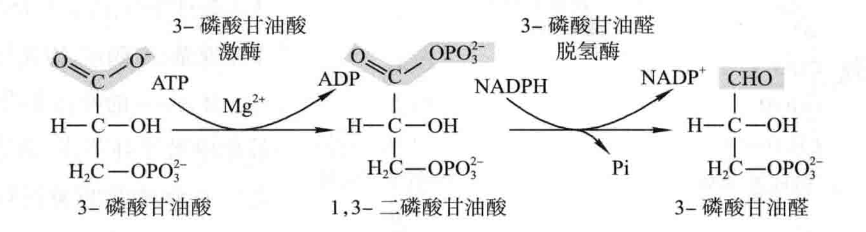


图 26-17 高等植物 Rubisco 的三维结构模型



该阶段类似于反方向的磷酸戊糖途径（见第二十四章“磷酸戊糖途径”），因此有人将卡尔文循环称为还原性的磷酸戊糖途径（reductive PPP）。

4. 卡尔文循环的总反应

综合 3 个阶段的反应，显然通过卡尔文循环净产生 1 分子 3- 磷酸甘油醛，至少需要有 3 分子 CO_2 进入循环，经过固定和还原以后共产生 6 分子 3- 磷酸甘油醛，其中 5 分子 3- 磷酸甘油醛进入再生阶段重新生成 3 分子 RUBP。在整个循环反应中，共消耗 6 分子 NADPH 和 9 分子 ATP，总反应式可以写成：

$$3CO_2 + 9ATP + 6NADPH + 6H^+ \longrightarrow 3\text{- 磷酸甘油醛} + 9ADP + 8Pi + 6NADP^+$$

e26-1 RUBP 的再生

如果以葡萄糖作为终产物，则总反应式可写成：

$$6CO_2 + 18ATP + 12NADPH + 6H^+ \longrightarrow \text{葡萄糖} + 18ADP + 18Pi + 12NADP^+$$

二、光呼吸

早在 1920 年，Otto Warburg 发现 C_3 植物叶内 O_2 相对量的提高能够抑制光合作用。究其原因，这与 Rubisco 本身的性质有关。已知 Rubisco 是一种双功能酶，它既有羧化酶的活性，又有加氧酶的活性，其活性中心与 CO_2 和 O_2 的 K_m 值分别是 9 μmol/L 和 350 μmol/L。

当 Rubisco 与 O_2 结合的时候，不仅不会导致 CO_2 的固定，反而会导致碳的流失。若 RUBP 与 O_2 结合，则先形成一个不稳定的中间产物，然后迅速裂解成 3- 磷酸甘油酸和磷酸乙醇酸（phosphoglycolate，PPG）。由于磷酸乙醇酸无法通过卡尔文循环代谢，而高浓度的磷酸乙醇酸对植物是有毒性的，因此植物细胞必须采取某种手段将它们及时处理掉。

Quiz5 有很多科学家企图通过蛋白质工程的手段改造 Rubisco，想在保留其羧化酶活性的同时使其丧失加氧酶，但都以失败告终，为什么？

植物处理 PPG 的方法为：在一种磷酸酶的催化下，磷酸乙醇酸被水解成乙醇酸；随后乙醇酸离开叶绿体进入过氧化物酶体，先被氧化成乙醛酸，然后被氨基化为 Gly；Gly 再进入线粒体，经甘氨酸脱羟酶和丝氨酸羟甲基转移酶的依次催化，变成 Ser，同时释放出 CO_2 和氨基；脱下来的氨基进入叶绿体，在消耗 ATP 的同时被转变为无毒的 Gln（图 26-18）。

由此可见，Rubisco 的加氧酶活性不仅导致了碳的流失，还浪费了 ATP，使光合作用的效率降低。植物体内发生的这种依赖于光、消耗 O_2、释放 CO_2 的过程称为光呼吸（photorespiration）。

虽然光呼吸与细胞的正常呼吸都是消耗 O_2、释放 CO_2 的过程，但是正常的细胞呼吸可以产生 ATP，而光呼吸却是浪费 ATP。

虽然 Rubisco 与 CO_2 的亲和力远大于 O_2，但大气中 O_2 的含量要比 CO_2 多得多，因此在 C_3 植物体内发生光呼吸是不可避免的。特别是在炎热干燥的气候条件下，光呼吸会变得更加严重，原因是在这种气候条件下，CO_2 的溶解性低于 O_2，植物为了保持水分，气孔经常处于关闭的状态，这使得大气中的 CO_2 进不去，而光反应中产生的 O_2 出不来。同时在强日照下，水的光解提速也会产生更多的 O_2。

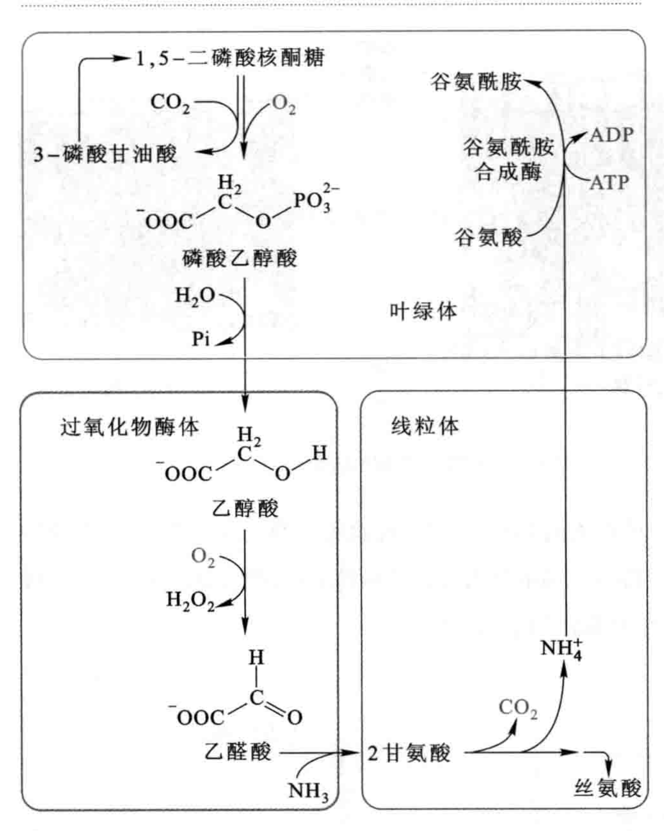


图 26-18 光呼吸的过程

Rubisco 所具有的这种缺陷被认为是进化史上的一个遗迹，也许在 Rubisco 刚刚出现的时候，根本就不存在这种问题。那时大气中游离的 O_2 甚少或者根本就没有，当时即使 O_2 能够与酶结合也很难和 CO_2 竞争。只是随着大气中 O_2 的含量逐渐升高，原来的缺陷便暴露出来了。事实上，在进化的历程中，Rubisco 也在完善之中，比如高等植物的 Rubisco 的羧化酶活性要比光合细菌的高出 8 倍之多。

据估计，光呼吸可导致 CO_2 固定的损失高达 50%，那是不是所有的植物对此就无计可施了呢？实际上，在长期的进化过程中，有两类植物，即 C_4 植物和景天酸代谢植物（crassulacean acid metabolism plant，CAM），它们已各自发展出一套独特的途径能将光呼吸降低到最低水平，下面就分别加以讨论。

三、C_4 循环与 C_4 植物

C_4 循环于 1966 年由 Hatch 和 Slack 在甘蔗和玉米中发现，因此又被称为 Hatch-Slack 途径。C_4 植物与 C_3 植物在叶的解剖学和代谢等环节具有一定的差别（表 26-4）。其克服光呼吸的方法并不是它的 RUBP 羧化酶无加氧酶的活性，而是它的叶片具有特殊的结构，再加上一种 PEP 羧化酶的出现，使得 Rubisco 周围的 CO_2 浓度提高 20～120 倍之多。

表 26-4 C_3 植物和 C_4 植物光呼吸的比较

类别	C_3 植物	C_4 植物
实例	小麦、水稻、棉花和苹果等	高粱、玉米、甘蔗和藜草等
光呼吸	较强	微弱
卡尔文循环	存在	存在
CO_2 的初级受体	1,5- 二磷酸核酮糖（RUBP）	磷酸烯醇式丙酮酸（PEP）
维管束鞘细胞是否含有叶绿体	无	有
羧化酶	Rubisco	叶肉细胞为 PEP 羧化酶；维管束鞘细胞为 Rubisco
羧化酶对 CO_2 的亲和性	中等	高
CO_2 固定的最初产物	3- 磷酸甘油酸（三碳中间产物）	草酰乙酸（四碳中间产物）
光合作用的细胞	叶肉细胞	叶肉细胞、维管束鞘细胞

与 C_3 植物不同的是，C_4 植物的叶脉被一层含有叶绿体的厚壁维管束鞘细胞（bundle-sheath cell）包围，而维管束鞘细胞则被薄壁叶肉细胞所包被（图 26-19）。C_4 植物的叶肉细胞并不表达 Rubisco，但表达 PEP 羧化酶；反之，它的维管束鞘细胞不表达 PEP 羧化酶，但表达 Rubisco。PEP 羧化酶与 Rubisco 最大的区别是没有加氧酶的活性，因此所催化的羧化反应即使在 CO_2 浓度很低的情况下也不会受到 O_2 的干扰。

C_4 循环的全过程是（图 26-20）：大气中的 CO_2 首先进入叶肉细胞，在 PEP 羧化酶的催化下，被固

定成草酰乙酸；草酰乙酸随后被还原成苹果酸，苹果酸则离开叶肉细胞，进入含有 Rubisco 的维管束鞘细胞，在苹果酸酶的催化下，转变成丙酮酸并释放出 CO_2；CO_2 在 Rubisco 的催化下进入卡尔文循环，而丙酮酸回到叶肉细胞，由丙酮酸磷酸双激酶（pyruvate phosphate dikinase）催化，在消耗 ATP 的同时转变成 PEP，PEP 因此得以再生。

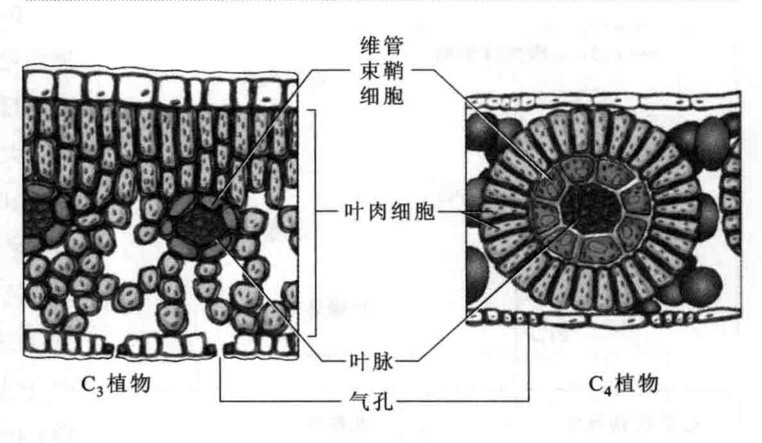


图 26-19　C_3 植物和 C_4 植物的叶片解剖结构

由此可见，C_4 循环在功能上相当于是一种耗能的 CO_2 充气泵，在这种气泵的作用下，进入叶肉细胞的 CO_2 被不断泵入维管束鞘细胞。这大大提高了鞘细胞内 CO_2 的局部浓度，使得 Rubisco 能够以最高的速率催化 CO_2 的固定，而其加氧酶的活性被降低到最低水平。

图 26-20　C_4 循环的全过程

由于 C_4 植物的光呼吸现象较弱，其光合作用的效率要比 C_3 植物高出 2 倍以上。需要注意的是，只有在炎热干燥的气候条件下，C_4 植物光合作用的效率才会高于 C_3 植物。在其他气候条件下，它反而要低于 C_3 植物，原因是 C_4 循环中 PEP 的再生消耗了大量的 ATP。

第五节　CAM 植物

CAM 植物多呈肉质，通常生活在半干旱和炎热的沙漠地带，例如仙人掌、菠萝和兰花。为了防止体内有限的水分发生不必要的流失，这一类植物在白天紧闭气孔，晚上气孔才开放。由于 CO_2 需要通过气孔进入体内，因此 CO_2 的固定只能在晚上进行，但光合作用只能在白天进行，包括光反应和暗反应。为了解决上述矛盾，在晚上气孔开放的时候，进入叶肉细胞的 CO_2 在 PEP 羧化酶的

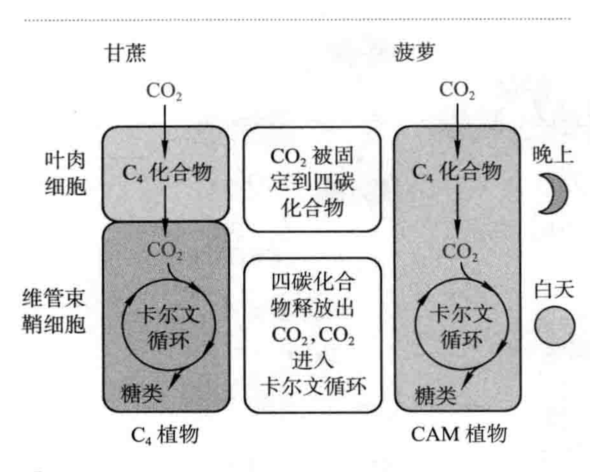


图 26–21 C_4 植物与 CAM 植物光呼吸的异同

催化下（这时 Rubisco 无活性，参看暗反应的调节）被暂且固定在草酰乙酸分子中。草酰乙酸随后转变为苹果酸，苹果酸则进入液泡并贮存在里面。一旦天亮，光反应开始产生 ATP 和 NADPH，Rubisco 也被激活，这时苹果酸离开液泡进入细胞质基质，在苹果酸酶的催化下一分为二，为卡尔文循环输送 CO_2。显然 CAM 植物以这种方式来固定 CO_2，在降低光呼吸方面与 C_4 植物具有异曲同工之妙（图 26–21）。

第六节　卡尔文循环的调节

构成暗反应的卡尔文循环与其他任何一条代谢途径一样，也受到严格的调控。调控的方式有光激活（light activation）和暗抑制（dark inhibition），那么光是如何激活卡尔文循环的呢？概括起来有以下几种方式：

（1）光诱导的 pH 变化（light-induced pH change）　光反应可导致质子被泵入类囊体腔，使得基质的 pH 升高到 8 左右，碱性的 pH 可激活基质内参与卡尔文循环的 Rubisco、1,6– 二磷酸果糖磷酸酶和二磷酸景天庚酮糖磷酸酶。

（2）光诱导的还原能力的产生（light-induced generation of reducing power）　卡尔文循环中的 5– 磷酸核酮糖激酶、1,6– 二磷酸果糖磷酸酶、1,7– 二磷酸景天庚酮糖磷酸酶和 3– 磷酸甘油醛脱氢酶都是巯基酶。当这些酶的巯基被氧化成二硫键以后，酶将丧失活性，光反应中产生的还原型铁氧还蛋白则通过硫氧还蛋白（thioredoxin）维持这些巯基酶处于有活性的状态（图 26–22）。

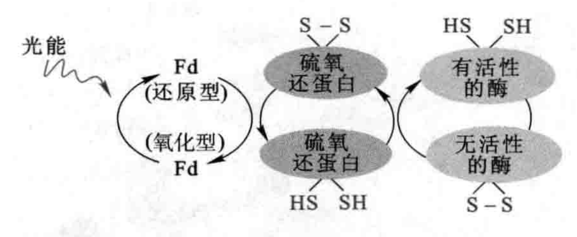


图 26–22　硫氧还蛋白对巯基酶活性的维持作用

某些植物在晚上能够合成一种天然的过渡态类似物抑制剂——1– 磷酸羧化阿拉伯糖醇（carboxyarabinitol-1-phosphate，CA1P），与 Rubisco 的活性中心结合，从而抑制 Rubisco 的活性。而解除抑制则需要 Rubisco 激活酶（Rubisco activase）的帮助，该激活酶能够促进 CA1P 离开 Rubisco 的活性中心，但它本身仍需要硫氧还蛋白的激活。

（3）光诱导的 Mg^{2+} 从类囊体流入基质　Mg^{2+} 对 Rubisco 的影响甚大。Rubisco 具有几种不同的状态：一种是无 Mg^{2+} 结合，也无氨甲酰化的状态。在这种状态下，酶采取的是封闭的构象，其活性中心紧密结合着 RUBP 或 CA1P，或者其他磷酸单糖，水分子被排除在外。因此，Rubisco 没有任何活性。第二种状态是无 Mg^{2+} 结合，但其 Lys 残基上的 ε– 氨基已被修饰成氨甲酰基，这时酶采取的是开放的构象，结合在活性中心的 RUBP 或其他磷酸单糖已被释放，但是酶仍然无活性。第三种状态是有 Mg^{2+} 结合的氨甲酰化形式，只有这种状态的 Rubisco 才具有活性（图 26–23）。

Quiz6 RUBP 既是 Rubisco 的底物，又是它的抑制剂。你认为这句话对吗？

Rubisco 的激活受到激活酶和 Mg^{2+} 的调节。在光照的条件下，Mg^{2+} 涌入基质，ATP 和 NADPH 被合成。这时候，环绕在 Rubisco 周围的具有 ATP 酶活性的激活酶被激活，在水解 ATP 的同时改变了 Rubisco 的构象，促进与 Rubisco 活性中心结合的 RUBP 或其他物质的释放。随后在 Mg^{2+} 的存在下，Rubisco 上一个特定的 Lys 残基经历甲酰化修饰。修饰后的 Rubisco 再结合一个 Mg^{2+} 就可以转变成活性形式。

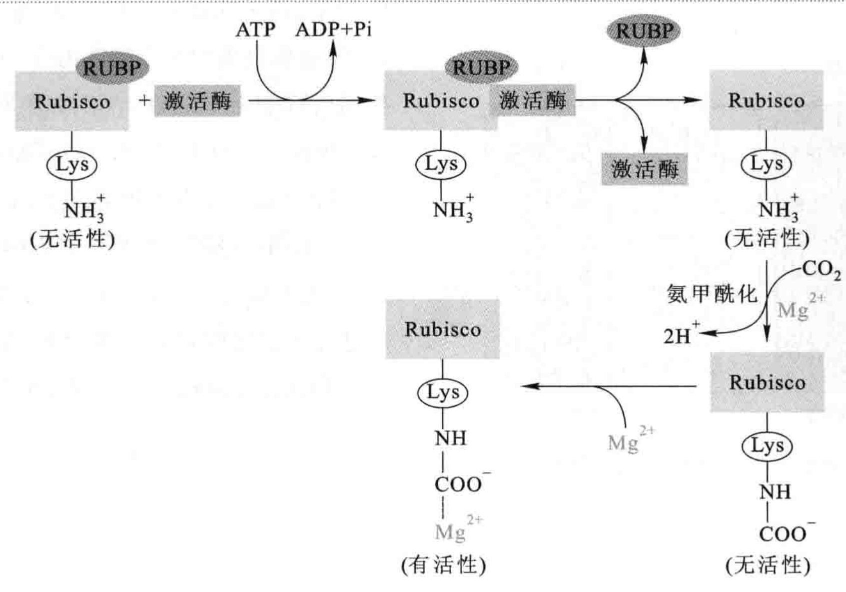


图 26-23 Mg^{2+} 和激活酶对 Rubisco 的调节

科学故事——卡尔文循环的发现

卡尔文循环的阐明既要归功于实验设计的巧妙，还要归功于电泳技术和同位素示踪技术的发展。20 世纪 50 年代，Melvin Calvin 及其同事开始使用单细胞光合有机体——小球藻（*Chlorella*）的悬液研究光合作用的暗反应。他们首先给小球藻细胞提供持续的光照和 CO_2，使光合作用处于稳态；然后他们在短时间内加入放射性同位素标记的 $^{14}CO_2$ 以标记循环的中间产物；接着将细胞悬液迅速倾入煮沸的乙醇溶液中杀死细胞，致使酶失活；最后使用双向纸电泳和放射自显影；分离、分析循环中的中间产物。

Calvin 通过这种方法，鉴定出碳同化过程中形成的 3- 磷酸甘油酸（PGA）和其他一些磷酸糖中间产物（图 26-24）。为了找出最早形成的碳同化中间产物，他们在细胞接触到 $^{14}CO_2$ 以后不同的时段（间隔 3 s、5 s、10 s 和 15 s）内杀死细胞，比较最后的电泳和放射自显影图谱，结果发现 3- 磷酸甘油酸是一个稳定的中间产物。于是他们认为，3- 磷酸甘油是最早形成的同化中间产物，而其他被标记的磷酸糖中间产物则是它的衍生物。

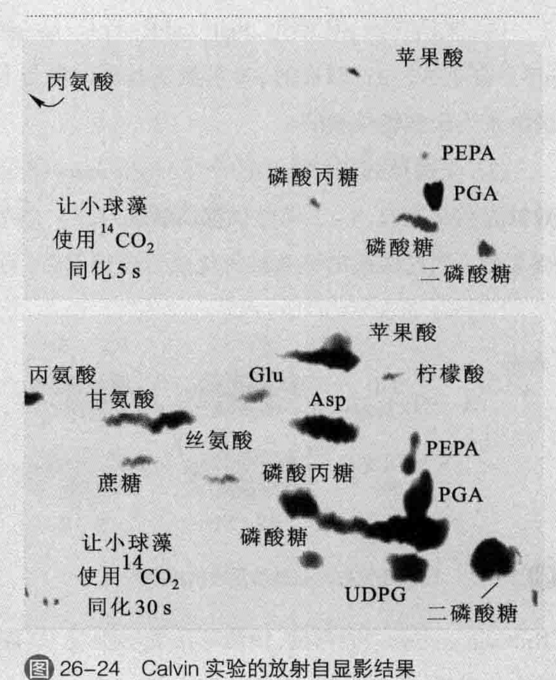


图 26-24 Calvin 实验的放射自显影结果

为了推导出整个同化途径，需要确定 ^{14}C 在每一个被标记的糖上的位置。Calvin 及其同事在小球藻短暂与 $^{14}CO_2$ 接触以后杀死细胞，分离出各中间产物，然后使用化学降解的方法，确定每一种中间产物上每一个位置 C 原子的 ^{14}C 的相对含量。结果显示，^{14}C 起初主要出现在 3- 磷酸甘油酸的羧基上。这个结果似乎在说明，CO_2 的起始受体是一个二碳化合物。这样的推论促使卡尔文拼命去寻找这种二碳中间产物，但却一无所获。

后来卡尔文发现，磷酸戊糖和二磷酸戊糖（核酮糖）也参与循环。这就意味着还有另外一种可能性，即五碳化合物作为 CO_2 的起始受体。这种在观念上的突破让他们很快确定了 1，5- 二磷酸核酮糖是 CO_2 的最初受体，并最终导致全部循环步骤的阐明。

卡尔文的工作大大促进了人们对地球上所有生命都很重要的光合作用暗反应的认识，他当之无愧地荣获了 1961 年的诺贝尔化学奖。



推荐网址：

1. http://en.wikipedia.org/wiki/Photosynthesis（维基百科有关光合作用的内容）
2. http://www.estrellamountain.edu/faculty/farabee/biobk/BioBookPS.html
3. http://web.virginia.edu/Heidi/chapter22/chp22frameset.htm（美国弗吉尼亚大学生物化学网络课程，有关光合作用的内容）
4. http://www.uic.edu/classes/bios/bios100/lecturesf04am/lect10.htm（美国伊利诺伊大学芝加哥分校提供的光合作用的内容）
5. http://www.wiley.com/college/boyer/0470003790/animations/photosynthesis/photosynthesis.htm（光合作用的动画）

参考文献：

1. Kohzuma K, et al. Thioredoxin-insensitive plastid ATP synthase that performs moonlighting functions. *Proc Natl Acad Sci* Early Edition, 1–6, 2012.
2. Umena Y, et al. Crystal structure of oxygen-evolving photosystem Ⅱ at a resolution of 1.9A°. *Nature*, 473: 55–61, 2011.
3. Scholes G D, et al. Lessons from nature about solar light harvesting. *Nature Chemistry*, 3(10): 763–774, 2011.
4. Fleming G R, et al. Architecture of a charge-transfer state regulating light harvesting in a plant antenna protein. *Science*, 320(5877): 794–797, 2008.
5. Arnon D I. The light reactions of photosynthesis. *Proc Nat Acad Sci*, 68: 2 883–2 892, 1971.
6. Kamp B M, et al. Two photochemical systems in photosynthesis. *Nature*, 190: 510–511, 1961.

第二十七章　糖原代谢

糖原又名动物淀粉，是糖类在动物体内的储存形式，主要存在于肝、肌肉和肾中，它在动物体内的存在有利于维持血糖浓度的稳定，同时可提供一种容易快速动员的短期储备燃料。一些细菌（如农杆菌）和真菌（如酵母）也能合成糖原。有关糖原在体内的分解和合成，特别是其调控机制，一直是人们研究的重点，已发现许多疾病与糖原的代谢失调有关。

本章将重点介绍糖原在高等动物体内的分解、合成及其调节机制。

第一节　糖原的分解

参与糖原分解的主要酶是糖原磷酸化酶（glycogen phosphorylase），它催化糖原的磷酸解反应，被磷酸解的葡糖残基总是位于非还原端（图 27–1）：

图 27–1　糖原磷酸化反应

此反应的 $\Delta G^{\ominus\prime}$ 为 3.1 kJ/mol，似乎对磷酸解不利，但由于细胞内无机磷酸的浓度远远高于 1–磷酸葡糖，因此反应几乎只能朝糖原分解的方向进行。磷酸解反应与水解反应十分相似，只不过是由无机磷酸代替水分子。

糖原磷酸化酶由两个完全相同的亚基组成，每一个亚基各共价结合一分子磷酸吡哆醛，结合的方式为希夫碱式。现已明确磷酸吡哆醛在磷酸化反应中的作用。参与磷酸化反应的只是与吡哆醛结合的磷酸基团，而与吡哆醛无关。

Quiz1 如何设计一个实验，证明糖原磷酸化酶的催化作用与吡哆醛基团无关？

糖原磷酸化酶催化的反应历程（图 27–2）是：在反应开始之前，作为底物的无机磷酸正好位于磷酸吡哆醛的磷酸基团与糖原底物之间，这种前后排列使得吡哆醛上的磷酸基团作为广义的酸碱催化剂发挥作用。反应开始时，作为底物的磷酸提供一个质子给非还原端要离开的葡糖残基 C4 上的氧原子，形成半椅式氧正离子（oxonium）中间物，同时它又从吡哆醛上的磷酸基团得到一个质子。随后，底物磷酸与半椅式中间物反应生成 1–磷酸葡糖，同时把质子还给吡哆醛上的磷酸基团。在整个反应过程中，水分子被完全排除在活性中心之外，以阻止水解反应的发生，很好地避免了 ATP 的浪费。

然而，糖原磷酸化酶的作用受到一定的限制：首先，它只能作用 α–1，4 糖苷键，不能作用于分支点上的 α–1，6 糖苷键；其次，它并不能作用于所有的 α–1，4 糖苷键，当遇到与分支点相距 4 个葡糖

残基的 α-1,4 糖苷键时就无法发挥作用，这时就需要脱支酶（debranching enzyme）。脱支酶是一种双功能酶，它的一种功能是具有 1,4→1,4- 葡萄糖基转移酶（glycosyl transferase）活性，借助此活性，可以将与分支点葡萄糖残基相连的不能再被磷酸解的 3 个葡萄糖单位，同时转移到邻近寡糖链的非还原端，并维持以 α-1,4 糖苷键连接。被转移到新位点上的葡萄糖残基可以正常地进行磷酸解，而遗留在分支点的葡萄糖残基，在脱支酶的第二种功能即 α-1,6- 糖苷酶（α-1,6-glucosidase）的活性作用下，被水解成游离的葡萄糖分子（图 27-3）。但除非机体在长时间饥饿的状态下，细胞内的一个糖原分子很少会被完全降解，通常只有它的外围部分降解，而留下一个核心，作为糖原再合成时的引物（参看糖原的合成）。

Quiz2 淀粉酶能不能作用于糖原？假定能够作用，那么每释放一个葡萄糖单位，经糖酵解可产生多少分子的 ATP？

糖原在磷酸解的时候，所有的非还原端可同时作为磷酸化酶的作用靶标，因此，糖原可以在很短的时间内动员。分支点越多，非还原端就越多，动员的速率就越快。

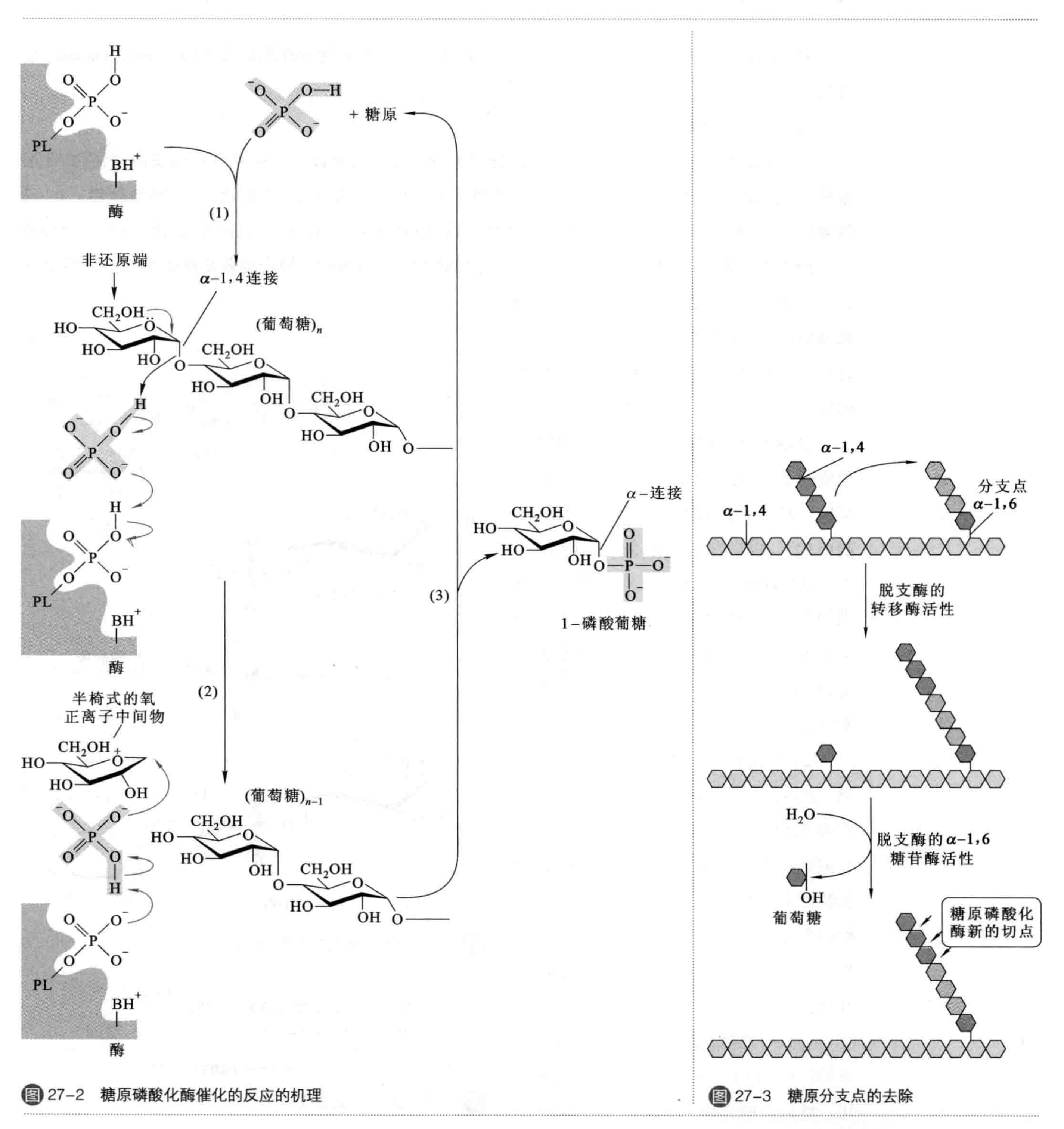


图 27-2　糖原磷酸化酶催化的反应的机理

图 27-3　糖原分支点的去除

磷酸解下来的 1- 磷酸葡糖经变位反应转变成 6- 磷酸葡糖，后者既可以直接进入糖酵解，又可以进入内质网，被参与糖异生最后一步反应的酶（6- 磷酸葡糖磷酸酶）水解成葡萄糖，并最终进入血液成为血糖的一部分。但由于肌细胞不表达 6- 磷酸葡糖磷酸酶，肌糖原磷酸解下来的 1- 磷酸葡糖只能进入糖酵解氧化，而不会脱掉磷酸根，更不会离开肌细胞成为血糖的一部分。显然，如果糖原磷酸解释放出的 1 个 1- 磷酸葡糖直接进入糖酵解，产生的 ATP 就要比游离的葡萄糖在糖酵解中产生的 ATP 多 1 个。这似乎意味着糖原的产能效率比游离的葡萄糖要高，其实不然，糖原是由游离的葡萄糖分子合成的，而一个游离的葡萄糖分子要变成糖原中的一个葡萄糖残基，需消耗 2 个 ATP 分子。

第二节　糖原的合成

糖原的合成并不是其磷酸解的逆反应，它需要由一种全新的酶即糖原合酶（glycogen synthase）来催化。

糖原的合成具有以下几个特征：

（1）需要活化的葡萄糖单位　游离的葡萄糖分子并不能直接作为糖原合酶的底物，而需要在细胞内首先被激活成 6- 磷酸葡糖，然后在磷酸葡糖变位酶的催化下，异构化成 1- 磷酸葡糖。1- 磷酸葡糖在动物和酵母体内受 UDPGlc 焦磷酸化酶（UPD-glucose pyrohorylase）的催化，转变成葡萄糖的一种活化形式——UDPGlc（图 27-4）。在细菌体内，1- 磷酸葡糖转变为葡萄糖的另外一种活化形式——ADPGlc。葡萄糖转变成 1- 磷酸葡糖的反应在糖酵解一章有专门的描述，这里只介绍 1- 磷酸葡糖转变成 UDPGlc 的反应。

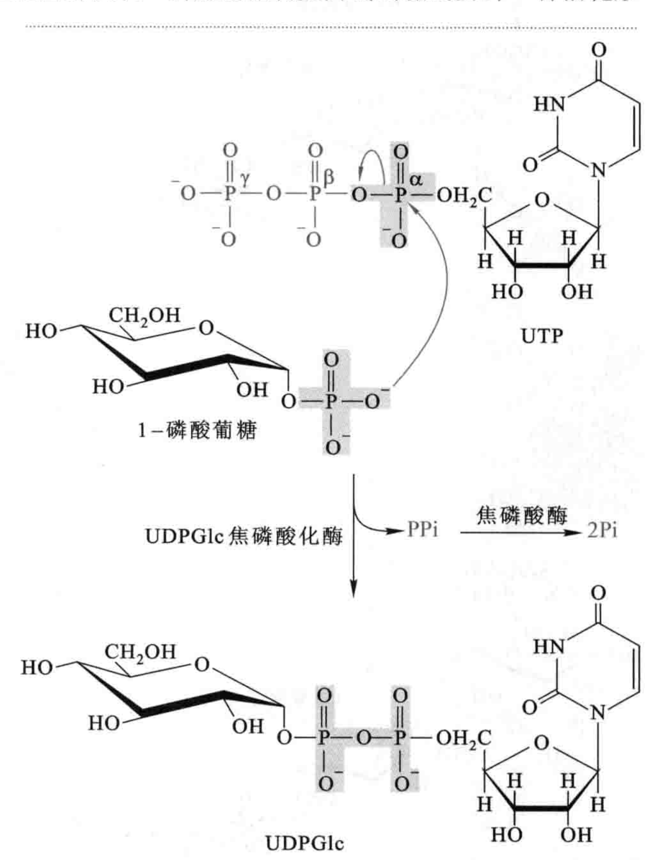


图 27-4　活化的葡萄糖单位的形成

单独由 1- 磷酸葡糖转变为 UDPGlc 反应的 $\Delta G^{\ominus\prime}$ 约为 0，这是一个典型的可逆反应。但由于生成的 PPi 在细胞内焦磷酸酶催化下，极易发生水解，致使总的 $\Delta G^{\ominus\prime}$ 为 -33.5 kJ/mol，因此总反应是不可逆的（图 27-5）。这种由焦磷酸水解驱动的反应会在很多代谢途径中出现，例如脂肪酸活化，氨基酸活化，DNA 复制和转录等（参看有关章节）。

反应	$\Delta G^{\ominus\prime}/(kJ\cdot mol^{-1})$
1- 磷酸葡糖 + UTP $\rightleftharpoons$ UDPGlc + PPi	~0
H_2O + PPi $\longrightarrow$ 2Pi	-33.5
总反应：1- 磷酸葡糖 + UTP $\longrightarrow$ UDPGlc + 2Pi	-33.5

图 27-5　UDPGlc 的合成

（2）需要引物　糖原合酶并不能催化糖原的从头合成（de novo synthesis），只能将葡萄糖单位从 UDPGlc 转移到事先已合成好的引物分子上。能充当糖原合成引物的分子有两种：一种是还没有完全降解的糖原分子，其残留的非还原端可直接作为引物；另外一种是蛋白质。对于酵母和动物而言，充当引物的蛋白质是糖原素（glycogenin）或糖原蛋白，它由两个相同的亚基组成，在糖原的“从头合成”中发挥作用。糖原素的作用十分重要，如果一个细

胞本来没有任何糖原分子，倘若再没有糖原素的话，即使它富含所有参与糖原合成的酶，也无法合成出一分子的糖原。

糖原素在糖原合成中的作用如下（图 27-6）：首先在它自带的酪氨酸葡糖基转移酶（tyrosine glucosyltransferase）的催化下，第一个葡萄糖单位从 UDPGlc 转移到糖原素 Tyr194 残基的羟基上，形成第一个 O 型糖苷键。随后，糖原素两个亚基之间相互催化（类似于 RTK 二聚体相互催化的自我磷酸化反应，参看第十七章“激素及其受体介导的信号转导”），将第二个葡萄糖单位转移到第一个葡萄糖单位的 4 号位羟基上，形成第一个 α-1,4 糖苷键。这样的反应可持续下去，直到形成一个七糖单位，之后由糖原合酶取而代之，但糖原素并没有解离。糖原合酶的活性需要糖原素的存在以及它与糖原素的直接接触。当糖原分子合成到一定大小以后，糖原合酶与糖原素脱离接触，糖原停止合成。因此，一个细胞内糖原分子的多少取决于糖原素分子的数目，而每一个糖原分子最多能合成到多大则取决于糖原合酶和糖原素何时脱离接触。

图 27-6 由糖原素引发的糖原合成

在能合成糖原的细菌体内还没有发现糖原素，它们用来合成糖原的引物可能就是糖原合酶本身。

(3) 合成方向是从还原端到非还原端　糖原链的延伸具有极性，只能从还原端向非还原端的方向进行，即新添加的葡萄糖单位只能加在非还原端。

(4) 分支需要分支酶（branching enzyme）　糖原合酶只能催化 α-1,4 糖苷键的形成，支链的形成需要分支酶。分支酶能将一个七糖单位，从一段长于 11 个葡萄糖残基的糖链非还原端，转移到邻近的糖链上，并以 α-1,6 糖苷键相连，新的分支点至少距离老的分支点 4 个葡萄糖残基以上（图 27-7）。

Quiz3 已在脑组织中发现有少量糖原的存在（约 0.5~1.5 g），它们主要存在于神经胶质细胞。这个发现有何意义？

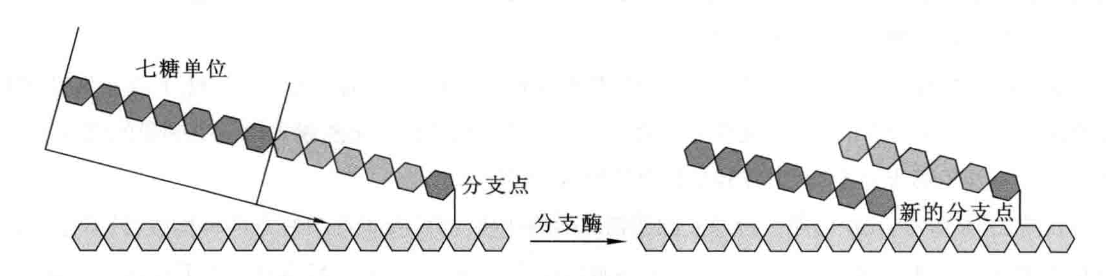


图 27-7 糖原分支的形成

框 27-1　生化点滴——三大能源贮备之 PK

机体有三大能源贮备——糖原、脂肪和蛋白质，这三种能源贮备各有利弊。

对于脂肪来说，其优点是贮能效率高，在细胞内的合成不会影响到细胞的渗透压，因此，在理论上它可以无限地贮存。但它的缺点是动员起来较慢，故只适合长期的能源贮备。另外，脂肪酸的氧化必须在有氧的条件下才能进行，因而，如果肌细胞使用脂肪作为唯一的储能物质，那当细胞出现氧债的时候，就会面临能源的短缺。由此可见，当机体在做并不激烈的运动的时候（如慢跑和长跑），脂肪才是一种不错的能源。

对于糖原来说，其优点有三点：首先，糖原动员起来很容易。糖原是高度分支的分子，其磷酸解反应可以在各非还原端同时展开，故当一种动物需要在短时间内消耗较多能量的时候（如做剧烈运动），糖原无疑是最好的能源。有的中跑和自行车运动员在比赛之前，会摄入大量的糖类，这样做的目的就是让机体贮备尽可能多的糖原供运动中使用。其次，糖原分解以及随后的糖酵解在有氧和无氧条件下都可以进行。选择糖原的最后一个好处是动物体内脂肪（其中的甘油和奇数脂肪酸除外）无法转化为葡萄糖，但红细胞、脑细胞和其他神经组织严重依赖葡萄糖的氧化供能。当饥饿的时候，肝糖原可迅速分解并转化为血糖，为脑组织等提供燃料。使用糖原的缺点是，贮能的效率要低于脂肪，在细胞内合成的时候，会影响细胞的渗透压，因此在细胞内积累有限，只适合充当短期能源贮备。人在饥饿一天一夜以后，糖原即被耗尽，无法供能。

对于蛋白质而言，它是机体最后的能源贮备，一般不会轻易动用。只有当机体长时间处于饥饿的状态时，蛋白质才会分解。分解产生的氨基酸主要是提供糖异生的原料，用来稳定血糖，并为脑细胞供能。蛋白质作为能源最大的缺点是，释放出来的氨基酸在分解的时候，会产生有毒的氨，机体需要对其解毒。而对氨的解毒耗能又耗水源。

第三节　糖原代谢的调节

如前所述，在相同的生理条件下，糖原分解和糖原合成都是放能反应，而且都发生在细胞质基质，显然，如果这两种途径同时发生，细胞将陷入无效循环。

糖原分解和合成的限速酶分别是糖原磷酸化酶和糖原合酶，这两种酶自然也就成为糖原代谢调节的主要目标。调节这两种酶活性的方式主要有两种，一是别构调节，二是受激素控制的“可逆磷酸化”调节。

一、别构调节

糖原分解代谢中调节的靶标是糖原磷酸化酶。肝细胞和肌细胞中的糖原磷酸化酶不完全相同，虽然它们属于同工酶，但是两者在结构和调节的某些方面仍然有差别，下面将分别加以讨论。

两种糖原磷酸化酶在结构与调节机制上的差异，实际上与两种糖原的功能差异有关。肌糖原完全是为肌细胞自身能量的需要而存在的，是“利己的”，而肝糖原是“利他的”，其存在的目的完全是为其他组织细胞（如脑细胞）提供能源和维持血糖浓度稳定。记住这两个事实将有助于理解糖原调节机制。

（一）肌糖原磷酸化酶的别构调节

肌糖原磷酸化酶与肝糖原磷酸化酶一样，都有 a 和 b 两种形式，通常 a 具有活性，b 没有活性，两者的区别仅仅在于前者的 Ser14 被磷酸化修饰了。当肌细胞受到肾上腺素作用或电刺激的情况下，b 即转变为 a（详见第十七章“激素及其受体介导的信号转导”）。

如图 27-8 所示，无论是哪一种形式的肌糖原磷酸化酶，都具有 T 态和 R 态两种构象，只不过 a 型的 R 态多，b 型的 T 态多。值得注意的是，T 态磷酸化酶的活性中心被一段自身肽链形成的环所掩盖，因此活性较低。

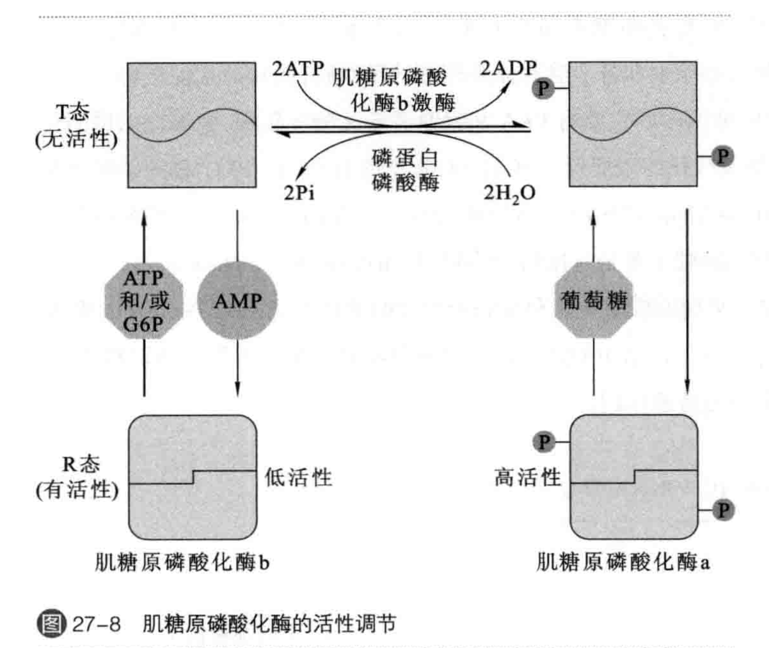


图 27-8 肌糖原磷酸化酶的活性调节

肌糖原磷酸化酶的别构效应物有 ATP、6- 磷酸葡糖和 AMP，它们只作用于 b 型磷酸化酶。其中 ATP 和 AMP 直接反映细胞的能量状态，6- 磷酸葡糖由 1- 磷酸葡糖变位而来，它的浓度实际上反映了糖原磷酸解进行的程度。

Quiz4 有人将糖原磷酸化酶的晶体与 AMP 保温，结果发现晶体破碎。对此你如何解释？

酶分子上有专门结合核苷酸的位点，当肌细胞内的能量较高时，高浓度的 ATP 与处于 R 态的 b 型磷酸化酶结合，促使它转变为无活性的 T 态，糖原磷酸解受阻；反之，当肌细胞内的能量较低时，高浓度的 AMP 与处于 T 态的 b 型磷酸化酶结合，促使它转变为有活性的 R 态，糖原分解随即启动。当细胞内糖原磷酸解到一定程度以后，6- 磷酸葡糖升高，细胞的能量提高，这时 6- 磷酸葡糖可以单独或者与 ATP 一起作用于处于 R 态的 b 型磷酸酶，促使 R 态转变为 T 态。

在大多数生理状态下，由于 ATP 和 6- 磷酸葡糖的抑制，肌细胞内的 b 型磷酸化酶无任何活性。然而，不论细胞内的 ATP、AMP 和 6- 磷酸葡糖的浓度如何变化，a 型磷酸化酶一直保持活性。

在肌细胞处于静息状态下，细胞内所有的磷酸化酶几乎都是无活性的 b 型（T 态）。当肌细胞收缩的时候，ATP 迅速消耗，AMP 浓度升高，T 态被激活成 R 态。与此同时，肌肉的运动导致肾上腺素释放，这有利于 b 型变成 a 型。由于肌肉细胞缺乏 6- 磷酸葡糖磷酸酶，6- 磷酸葡糖不会水解成葡萄糖，而只能留在肌细胞内进行氧化放能。

Quiz5 肌肉细胞有哪几种防止能量缺乏的措施？

（二）肝糖原磷酸化酶的别构调节

人类肝糖原磷酸化酶与肌糖原磷酸化酶的 90% 氨基酸序列是相同的，但是，正是那 10% 的差异使得两者在调控机制上存在一些差别。

首先，肝细胞内是 a 型磷酸化酶对别构效应物产生强烈的反应，而不是 b 型磷酸化酶；其次，别构效应物并不是 AMP、ATP 和 6- 磷酸葡糖，而是葡萄糖；此外，别构效应物的作用方式也不一样。

既然肝糖原磷酸化酶的主要功能是为其他组织提供葡萄糖，那么，当机体在某一个时刻能从其他途径获得葡萄糖，如饱食一顿富含糖的食物以后，血糖浓度就会迅速提高，此时血液中多余的葡萄糖可通过肝细胞膜上高效的 GLUT2 进入肝细胞，作为别构效应物与 a 型磷酸化酶结合，促使它变成无活性的 b 型磷酸化酶。至于葡萄糖如何使磷酸化酶从 a 型转变为 b 型，将在后面加以讨论。

（三）糖原合酶的别构调节

与糖原磷酸化酶一样，糖原合酶也具有 a 型和 b 型，但与磷酸化酶不同的是，a 型主要是去磷酸化形式，b 型主要是磷酸化形式。

作为别构效应物直接影响糖原合酶活性的分子是 6- 磷酸葡糖，它能够降低 b 型合酶对 UDPGlc 的 K_m 值，从而使磷酸化的合酶也能够合成糖原。

二、受激素控制的“可逆磷酸化”调节

别构效应物对糖原磷酸化酶和糖原合酶的直接调节仅仅是在最低一级水平对糖原代谢进行调节，更精细、更敏感和更灵活的调节需要在高一级水平上进行，这必然涉及更多的生物分子，如激素、受体、G 蛋白、第二信使、蛋白质激酶和磷蛋白磷酸酶等。

Quiz6 肌糖原的代谢受不受胰高血糖素的调控？为什么？

参与糖原代谢调节的激素有胰岛素、胰高血糖素和肾上腺素，肾上腺素主要作用于肌细胞，胰高血糖素主要作用于肝细胞。其中，胰高血糖素和肾上腺素促进糖原的磷酸解，而胰岛素促进糖原的合成。参与调节糖原磷酸化酶磷酸化修饰的激酶主要有 PKA 和糖原磷酸化酶 b 激酶，被修饰的磷酸化酶位点是 Ser14，而参与调节糖原合酶磷酸化的激酶至少有 11 种，常见的有 PKA、PKC、糖原磷酸化酶 b 激酶、糖原合酶激酶（glycogen synthase kinase，GSK）、酪蛋白激酶（casein kinase，CK）Ⅰ和酪蛋白激酶Ⅱ，磷酸化位点至少有 9 个。磷蛋白磷酸酶则主要是磷蛋白磷酸酶-1（phosphoprotein phosphatase type 1，PP-1）。关于胰高血糖素和肾上腺素激活糖原磷酸化酶和抑制糖原合酶的机制见图 27-9，详细过程见第十七章"激素及其受体介导的信号转导"，这里主要进一步介绍糖原磷酸化酶 b 激酶的结构和功能，同时重点介绍 PP-1 在糖原代谢调节中独特的作用。

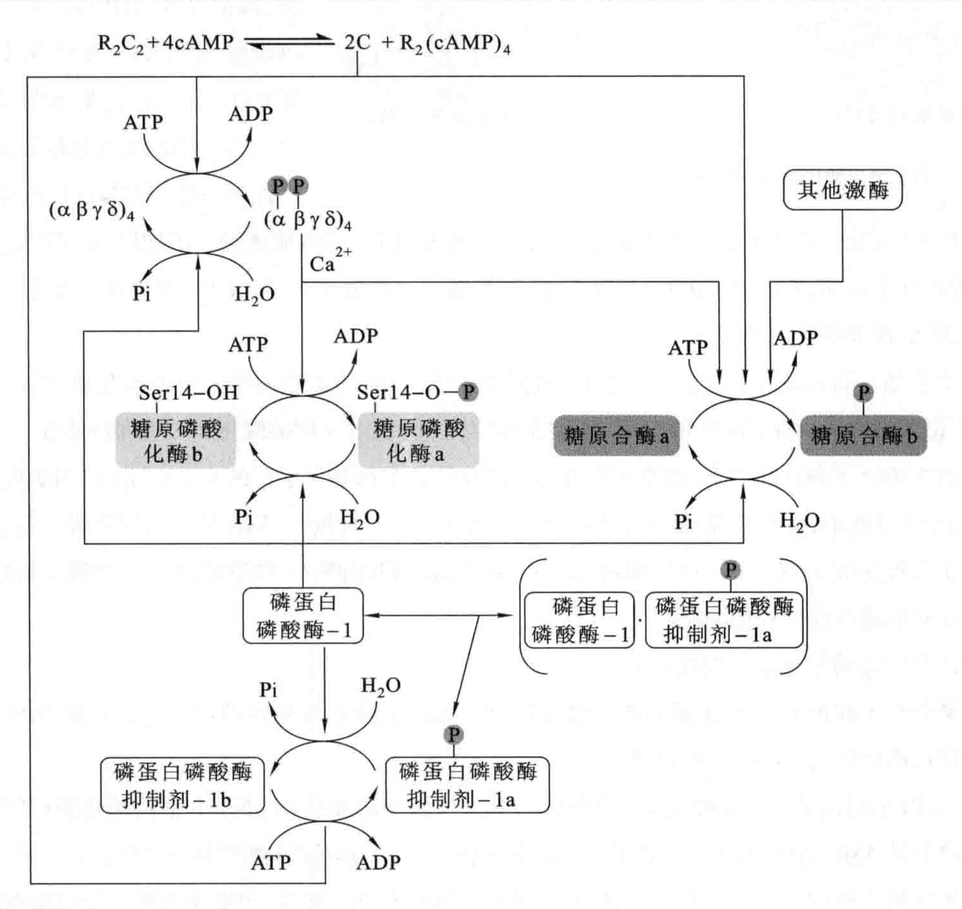


图 27-9 糖原代谢的调节

糖原磷酸化酶 b 激酶由 α、β、γ 和 δ 亚基组成，每一个亚基有 4 个拷贝，因此可简写成 $(\alpha\beta\gamma\delta)_4$。其中 β 亚基含有磷酸化位点，为催化亚基，δ 亚基本质上是钙调蛋白。糖原磷酸化酶 b 激酶的底物既有糖原磷酸化酶 b，又包括糖原合酶 a（图 27-10）。

当细胞内的 PKA 系统被肾上腺素或胰高血糖素激活以后，磷酸化酶 b 激酶的 β 亚基被磷酸化修饰，这时它只有部分活性，要想达到完全的活性，δ 亚基就必须结合有 Ca^{2+}；同样，如果仅仅 δ 亚基结合有 Ca^{2+}，而 β 亚基没有被磷酸化，也只有部分活性。因此，磷酸化酶 b 激酶的完全激活需要 Ca^{2+} 和 PKA 的共同作用。

细胞内的 Ca^{2+} 通常贮存在内质网，因此细胞质基质的 Ca^{2+} 浓度一般很低，难以结合钙调蛋白。因此，要想完全激活糖原磷酸化酶 b 激酶，Ca^{2+} 必须得从内质网里释放出来。

肝细胞和肌细胞促进 Ca^{2+} 释放的机制完全不同，前者通过肾上腺素与它的 α_1-受体结合，经过 Gq 蛋白激活 PLC-β，产生第二信使 IP_3 来完成反应（参看第十七章"激素及其受体介导的信号转导"）；后者直接由神经脉冲控制，肌肉收缩的同时，就有大量 Ca^{2+} 从肌浆网里面释放进入细胞质基质。

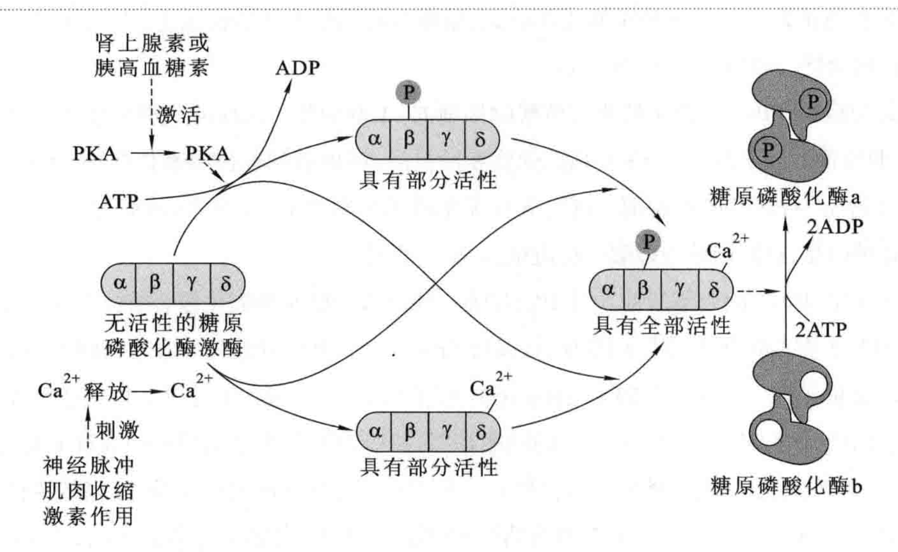


图 27-10 糖原磷酸化酶 b 激酶的活性调节

PP-1 是一种非常重要的磷蛋白磷酸酶，由催化亚基和糖原结合亚基（glycogen-binding subunit，简称 G 亚基）组成。G 亚基负责与糖原结合，而 PP-1 只有与糖原结合以后才会有活性。它的底物包括糖原磷酸化酶 a、糖原磷酸化酶 b 激酶、磷蛋白磷酸酶抑制剂 -1a 和糖原合酶 b。它的作用就是催化这几种蛋白质的去磷酸化反应，一方面抑制了糖原磷酸化酶 b 激酶、糖原磷酸化酶和磷蛋白磷酸酶抑制剂的活性，另一方面则刺激了糖原合酶的活性。而它本身的活性既受到磷蛋白磷酸酶抑制剂 -1（phosphoprotein phosphatase inhibitor 1，PPI-1）的直接调节，又受到蛋白质激酶的调节。能够修饰 PP-1 的蛋白质激酶有两种，一种是 PKA，另一种是受胰岛素激活的蛋白质激酶。

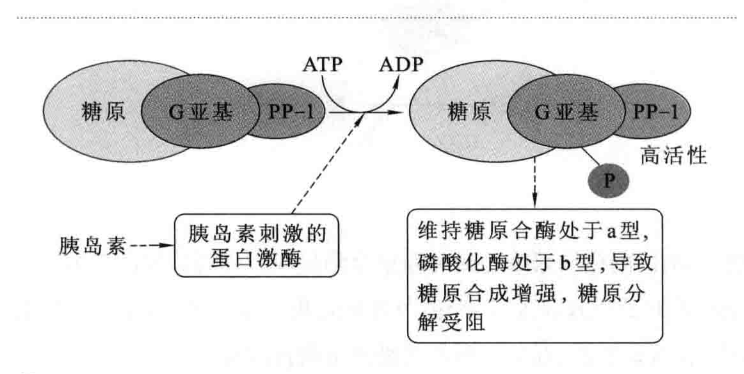


图 27-11 PP-1 对糖原代谢的调节

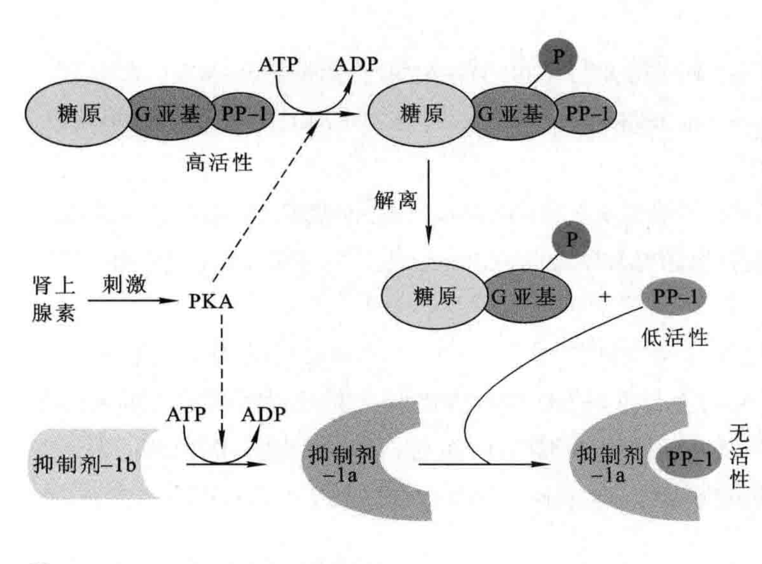


图 27-12 肾上腺素对糖原代谢的调节

如图 27-11 所示，PP-1 能够通过它的 G 亚基与糖原结合，当细胞受到胰岛素作用的时候，胞内一种特殊的蛋白质激酶被激活，PP-1 被磷酸化修饰，在这种状态下，PP-1 仍然与 G 亚基结合在一起，酶活性极高。高活性的 PP-1 可导致细胞内糖原磷酸化酶和糖原磷酸化酶 b 激酶都处于低活性的去磷酸化状态，而糖原合酶活性较高，此时细胞内的糖原合成增强；反之，当细胞受到胰高血糖素或肾上腺素作用的时候，细胞内的 PKA 被激活，导致 PP-1 发生磷酸化修饰，但其修饰位点不同于被胰岛素激活的蛋白质激酶修饰的位点，因此产生的效应是不一样的。被 PKA 修饰的 PP-1 与它的 G 亚基解离，其酶活性急剧下降（图 27-12）。在这种情况下，细胞内的糖原磷酸化酶和糖原

磷酸化酶 b 激酶将处于高活性的磷酸化状态，而糖原合酶则处于低活性状态，因此，此刻细胞内的糖原磷酸解作用增强，而糖原合成受到抑制。

在肌肉细胞中，PP-1 还受到磷蛋白磷酸酶抑制剂 -1 的调节。这种抑制剂本身也是蛋白质，也能受到 PKA 的修饰。当细胞内的 PKA 系统受到激活以后，该抑制剂也被磷酸化修饰，而从无活性的 b 变成有活性的 a。有活性的抑制剂 a 便与从 G 亚基解离出来的 PP-1 结合，从而导致 PP-1 的活性完全丧失。有趣的是，抑制剂的去磷酸化反应也由 PP-1 催化。

在肝细胞中，PP-1 的调节活性主要由它与哪一种形式的糖原磷酸化酶 a 结合而定，如前述，磷酸化酶有 R 态和 T 态两种构象，两种构象都可以与 PP-1 结合，但只有处于 T 态时，磷酸化酶上的 Ser14 才会暴露出来被水解。因此，当磷酸化酶 a 处于 R 态时，PP-1 即使结合上去，也不起作用，只有当它转变为 T 态以后，才会被水解掉 Ser14 上的磷酸基团，并变为无活性的 b，将磷酸化酶 a 变为无活性的磷酸化酶 b。在肝细胞内，调节糖原磷酸化酶 a 从 R 态向 T 态转变的物质是游离的葡萄糖分子，因此，在某种意义上，糖原磷酸化酶 a 可被视为葡萄糖在细胞内浓度变化的感应器。如图 27-13 所示，当肝细胞内游离的葡萄糖分子浓度较高时，它与处于 R 态的磷酸化酶 a 结合，促使 R 态转变为 T 态，而处于 T 态的磷酸化酶 a 立即被 PP-1 水解转变成无活性的磷酸化酶 b。与此同时，PP-1 从已变成 b 型的磷酸化酶上解离下来，以糖原合酶 b 为底物，使之变成有活性的糖原合酶 a。

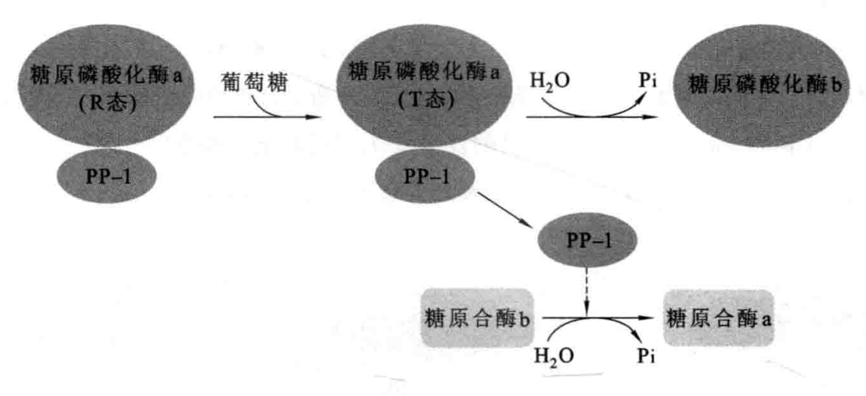


图 27-13 葡萄糖对肝糖原代谢的调节

总之，无论是肝细胞还是肌细胞，其胞内的糖原磷酸化酶和糖原合酶是同时受到调控的。哪一种被激活、哪一种被抑制完全取决于细胞当时的生理状况，整个调节过程是极为复杂的，既受到别构效应物的直接调节，又受到“可逆磷酸化”的级联控制，而后一种方式最终由激素控制。

科学故事——糖原素的发现

对糖原代谢的研究造就了许多著名的生化学家，其中就有 6 人因此而荣获诺贝尔生理学或医学奖。他们包括：Cori 夫妇、Luis Leloir、Sutherland、Edmond H. Fische 和 Edwin G. Kreb。这里介绍的是糖原合成代谢的研究促成糖原素发现的故事。

最初人们对糖原合成的认识很肤浅，而且很多观点并不正确。例如认为糖原合成仅仅是糖原磷酸解的逆反应。在体外，糖原的确可在糖原磷酸化酶和分支酶的催化下形成。有人发现该反应需要引物，但引物究竟是什么并不清楚。

Williarm J. Whelan 于 1950 年也加入到研究糖原合成代谢之中。但他开始研究的是土豆里面的淀粉磷酸化酶催化 1- 磷酸葡糖合成直链淀粉的反应。在此研究中，他发现麦芽五糖和六糖可作为淀粉合成的引物。然而，他很快发现，体内催化糖原合成的酶并非磷酸化酶，而是另外一种酶，即糖原合酶。对于糖原合酶来说，它催化的反应至少要在麦芽四糖（maltotetraose）这样大小的引物上才能进行。那么麦芽四糖又是怎样合成的呢？

1962年，Carl Cori及其同事报道，肌肉细胞的糖原磷酸化酶不需要引物就可以将1-磷酸葡糖合成为糖原。他们把这样的合成称为“从头合成”。但Whelan在重复Cori的实验中，发现Cori是错的。于是，有关糖原合成需要不需要引物以及引物是如何合成的问题引起了很大的争议和关注。

直到1975年，Leloir实验室的Clara Krisman发现，将大鼠肝细胞的抽取物与^{14}C标记的UDPGlc保温，一段时间以后，用三氯乙酸（TCA）沉淀，可得到类似糖原的物质。于是，Krisman认为，糖原连在一种蛋白质分子上，而这种蛋白质可能就是糖原合成真正的引物，而不是之前认为的游离麦芽四糖。Krisman还进一步推想，这种蛋白质充当糖原起始合酶（glycogen initiator synthase），由它催化麦芽四糖的形成。

当Whelan得知这个消息以后，决定邀请Krisman一起来寻找可能充当糖原合成引物的蛋白质。他们的思路是首先去找细胞内与糖原共价结合的蛋白质。尽管他们也意识到，作为引物的蛋白质与糖原的结合可能不是永久性的，很可能在麦芽四糖合成好以后就与糖原分离了，但他们仍然觉得值得一试。

关于细胞内的糖原与蛋白质结合这种观点并不新鲜，事实上早在1886年就有相关的报道了。到了1930年左右，Willstaette认为细胞内的糖原有两种，一种与蛋白质结合，另一种不与蛋白质结合，其中前一种可以被低浓度的TCA沉淀。但到了1961年，Stetten对两种可能的糖原组分进行了详细的研究，最后得出与蛋白质结合的糖原是实验假象的结论，并认为麦芽四糖是合成糖原的引物。于是，在此后很长的一段时间内，再没有人对与糖原结合的蛋白质这个问题感兴趣了。

现在，有关糖原中含有蛋白质的问题又被提出，并引起了Whelan等人浓厚的兴趣。他们很快就发现肝糖原的确含有蛋白质，但不久就把研究对象改为肌糖原，因为考虑到肌糖原含量更加丰富。到了1979年，他们的研究终于取得了突破性的进展，一名来自波兰的访问学者Joseph Lomako，从肌糖原中分离到一种M_r为3.7×10^4的蛋白质，这种蛋白质与葡萄糖残基共价结合，他们将这种蛋白质命名为糖原素。

在得到糖原素以后，他们决定先确定它与糖原之间的共价连接方式。但这一步研究并不顺利，主要有三个原因：一是两者的连接方式是全新的，与通常的蛋白质－糖类的连接方式差别很大；二是蛋白质分子中与糖原连接的Tyr残基在酸性水解条件下很容易被破坏；三是缺乏后续的研究经费。Whelan在已经发现兔子的肌糖原都有共价结合的蛋白质的时候，仍然没有能够得到NIH基金的后续支持，因为评审专家认为这项研究进展太慢。但Whelan并没有因此放弃对糖原素的研究，在没有相关经费的支持下，他的两个研究生Ignacio Rodriguez和Brent Kirkma继续研究了6年。到了1985年，他们终于确定了糖原素和糖原之间的连接方式。有了这个结果以后，他们想从NIH申请进一步研究的基金，以得到纯的糖原素，并确定它是否就是糖原从头合成的引物。然而，让他们意想不到的是，他们的基金申请再次被拒。被拒的理由竟然是，评审专家对重复认为糖原素是糖原合成引物的推论感到困扰。要知道没有经费的支持，任何一项研究都是寸步难行的。虽然这时Whelan已达到退休的年龄，但仍没有放弃对糖原素的研究。他继续用着一点点退休基金勉强支撑着下一步的研究。终于有一天，当他们将兔子肌肉细胞的抽取物、^{14}C标记的UDPGlc和含有Mn^{2+}的缓冲液混合在一起以后，发现有蛋白质被同位素标记。在使用SDS-聚丙烯酰胺凝胶电泳进行分离后，发现被标记的蛋白质位于3.7×10^4的位置。经糖原素抗体的Western印迹验证，确定被标记的蛋白质就是糖原素。于是，糖原素作为糖原从头合成的引物终于得到了确认。

推荐网址：

e27-1 本章小结填填看

1. http://themedicalbiochemistrypage.org/glycogen.html（完全免费的医学生物化学课程网站，有关糖原代谢的内容）
2. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb1/part2/glycogen.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关糖原代谢的内容）
3. http://web.virginia.edu/Heidi/chapter23/chp23.htm（美国弗吉利亚大学生物化学网络课程，有关糖原代谢的内容）

4. http://cmgm.stanford.edu/biochem200/glycogen/（美国斯坦福大学医学院提供的有关糖原代谢的内容）

参考文献：

1. Roach P J. Phosphate Incorporation during Glycogen Synthesis and Lafora Disease. *Cell Metabolism*, 13, 274–282, 2011.
2. Simoni R D, et al. Carbohydrate metabolism: glycogen phosphorylase and the work of Carl F. and Gerty T. Cori. 1928–1943. *J Biol Chem*, 277(29): 18e, 2002
3. Roach P J, et al. Crystal structure of the autocatalytic initiator of glycogen biosynthesis, glycogenin. *J Mol Biol*, 319: 463–477, 2002.
4. Barford D, et al. Structural mechanism for glycogen phosphorylase control by phosphorylation and AMP. *J Mol Biol*, 218: 233–260, 1991.

第二十八章　脂肪、磷脂和糖脂的代谢

脂肪、磷脂和糖脂这三种脂质的主要功能分别是充当能量贮存者、细胞膜的组分和细胞表面的信号分子，它们在机体内的代谢转变也一直受到人们的关注，例如，脂肪代谢与肥胖的形成有直接的关联，对它的研究将为人类深入了解肥胖的机理，开发有效的减肥药物，并最终彻底战胜肥胖提供理论指导。

本章将主要讨论脂肪和磷脂的代谢，而有关糖脂代谢的内容仅做简单介绍。

第一节　脂肪代谢

一、脂肪的水解

脂肪的水解是在脂肪酶催化下完成的，这里以高等动物脂肪组织中的脂肪为例，简要说明该组织中的脂肪是如何被水解的。

高等动物的脂肪主要以脂滴的形式贮存在脂肪细胞（见第十六章"脂质与生物膜"），其动员是一种受激素控制的过程。胰高血糖素和肾上腺素促进脂肪的水解，而胰岛素则抑制脂肪的水解，同时促进脂肪的合成。如图 28-1 所示，在饥饿或禁食的情况下，体内的胰高血糖素和肾上腺素开始分泌，随着血液循环到达脂肪细胞，与位于脂肪细胞膜上的 G 蛋白偶联受体结合，依次通过 Gs 蛋白、腺苷酸环化酶和 PKA（参看第十七章"激素及其受体介导的信号转导"），最终导致胞内的脂外被蛋白和激素敏感性脂肪酶（hormone-sensitive lipase，HSL）发生磷酸化修饰，脂外被蛋白因发生磷酸化而丧失活性，HSL 则由于磷酸化被激活。脂外被蛋白的失活导致脂滴表面的保护性屏障开放，允许 HSL 接触并分解脂滴内储存的脂肪分子，使其水解成二酰甘油和 FFA，而二酰甘油的进一步水解是由另一类不受激素控制即对激素不敏感的脂肪酶催化的。

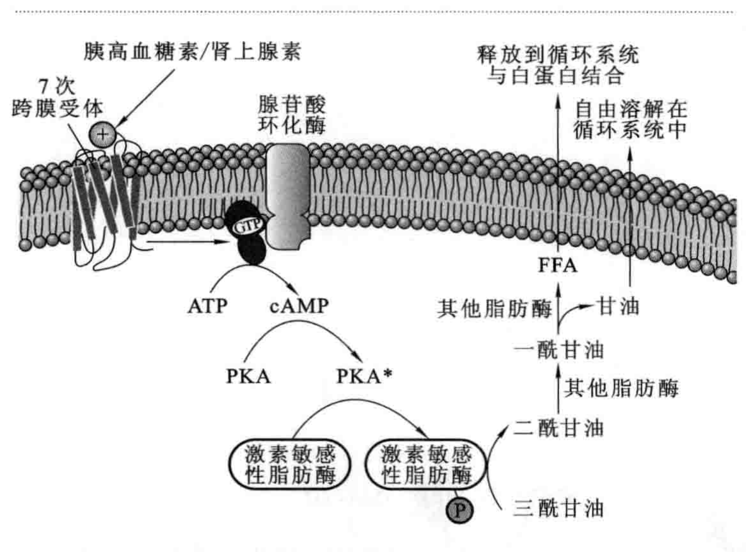


图 28-1　受激素控制的内源性脂肪动员（Berg 等，2006）

Quiz1　经常饮用绿茶和咖啡，特别是在饥饿的时候有利于减肥，为什么？

经水解释放出来的 FFA 通过自由扩散进入血流，并与血浆中的白蛋白结合。白蛋白的作用是帮助溶解性不高的中、长链脂肪酸在血液中的运输。一分子白蛋白可结合多个 FFA。在白蛋白的帮助下，FFA 被转移到骨骼肌、心肌和肝等组织细胞内进一步氧化分解（参看第二十九章"脂肪酸代谢"）。在进入细胞以后，长链的 FFA 与胞内特定的结合蛋白结合，因此细胞内真正游离的长链 FFA 很少。

甘油则可以直接溶解在水相中，通过循环系统进入肝或其他细胞被利用（参看第二十二章"糖酵解"和第二十五章"糖异生"）。

二、脂肪的合成

脂肪的合成实际上就是甘油的 3 个羟基被脂酰化的过程，然而在细胞内，游离的甘油和脂肪酸是难以缩合成酯的，在缩合之前两者各自都需要活化，其中甘油被活化成 3- 磷酸甘油，脂肪酸被活化成脂酰 -CoA（图 28-2）。

图 28-2　脂肪合成的序列反应

(1) 甘油的活化　活化甘油既可以由甘油直接激活而来,又可以由糖酵解的中间产物磷酸二羟丙酮还原而来,相关的反应参看第二十二章“糖酵解”。

(2) 脂肪酸的活化(参看第二十九章“脂肪酸代谢”)。

(3) 磷脂酸的形成　在转酰酶(transacylase)催化下,甘油磷酸上两个游离的羟基和两个脂酰 -CoA 先后缩合形成磷脂酸(phosphatidic acid)。

Quiz2 哪些物质的合成以磷脂酸为前体?

(4) 二酰甘油的形成　在磷脂酸磷酸酶的催化下,磷脂酸被水解为二酰甘油,从而暴露出第三个自由羟基。

(5) 脂肪的形成　在转酰酶的催化下,二酰甘油上的游离羟基被脂酰化成脂肪即三酰甘油。

框 28-1　生化与健康——TZD 治疗Ⅱ型糖尿病的机理

Ⅱ型糖尿病也称成人发病型糖尿病,多在 35 岁之后发病,占糖尿病患者 90% 以上。Ⅱ型糖尿病患者体内产生胰岛素的能力并非完全丧失,有的患者体内胰岛素甚至产生过多,但患者对胰岛素的作用有抗性。

肥胖被认为是诱发Ⅱ型糖尿病的主要原因,而有关Ⅱ型糖尿病患者体细胞对胰岛素的作用不敏感原因还不十分清楚。但目前市场上治疗Ⅱ型糖尿病的少数几种药物——噻唑烷二酮类(thiazolidinedione, TZD),不仅能让病人体内胰岛素发挥作用,还能增加体重。这是为什么呢?

原来,TZD 能够激活脂肪细胞产生甘油激酶,这种酶在脂肪细胞内一般不表达,它催化甘油变成甘油磷酸,这是脂肪合成中的一步反应。TZD 的作用等于是在脂肪细胞内引入一种“无效”循环:一方面,脂肪细胞分解脂肪,释放出游离的脂肪酸和甘油,另一方面,在脂肪细胞中,由 TZD 诱导产生的甘油激酶又将脂肪酸和甘油转变成脂肪贮存起来。由于甘油激酶催化脂肪合成的速率更快,因此,最终导致流入脂肪组织的脂肪酸更多,从而减少了循环系统中游离脂肪酸的含量,使病人对胰岛素的敏感性增强。

当机体需要燃料的时候,比如糖尿病患者的细胞不能吸收血糖时,脂肪细胞开始分解脂肪,产生甘油和 FFA。甘油可随循环进入肝细胞,在肝细胞内经糖异生变成葡萄糖,而脂肪酸可进入肌肉细胞氧化分解。TZD 的作用是逆转上述过程,它能与脂肪细胞膜上的 PPAR-γ 受体结合,诱导脂肪细胞产生甘油激酶,促进脂肪的重合成。这也让脂肪细胞能够“捕获”路过的脂肪酸,于是使血流中的 FFA 水平下降。

第二节　磷脂代谢

一、磷脂的分解

(一) 甘油磷脂的分解

水解甘油磷脂的酶统称为磷脂酶(phospholipase),它们存在于各种组织(与膜结合或以游离的形式存在)、胰液和毒液(如蛇毒和蜂毒等)之中。不同的磷脂酶对于甘油磷脂上酯键的特异性不同,图 28-3 显示了 4 类主要的磷脂酶对酯键作用的特异性和分布上的差别。

Quiz3 预测:磷脂酰丝氨酸被磷脂酶 A_1、A_2、C 和 D 水解后的产物分别是什么。

(二) 鞘磷脂的分解

鞘磷脂的水解是在溶酶体内进行的,参与水解的酶有鞘脂酶(sphingomyelinase)和神经酰胺酶(ceramidase)(图 28-4),其中鞘脂酶将鞘磷脂水解成神经酰胺和磷酸胆碱,神经酰胺酶则将神经酰胺进一步水解成神经鞘氨醇和 FFA。

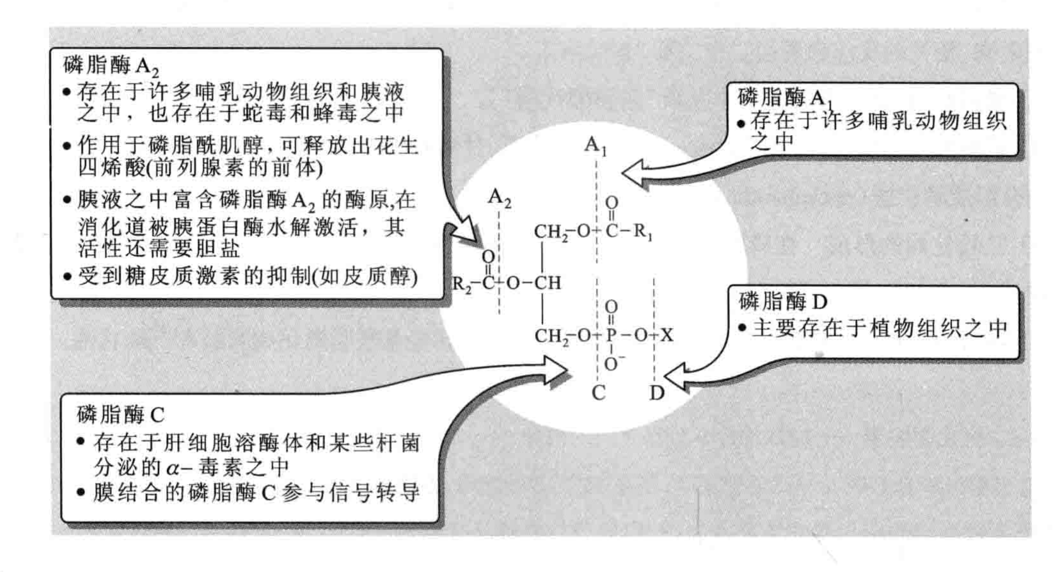


图 28-3　不同磷脂酶的性质比较

鞘脂酶

神经酰胺

$CH_3(CH_2)_{12}-CH=CH-\underset{OH}{\overset{H}{C}}-\underset{NH}{\overset{H}{C}}-CH_2-O-\underset{O^-}{\overset{O}{P}}-OCH_2CH_2\overset{+}{N}(CH_3)_3$

磷酸胆碱

神经酰胺酶

$CH_3(CH_2)_n-\underset{O}{C}$

脂酰基

图 28-4　鞘磷脂的水解

尼曼 – 皮克病(Niemann-Pick disease)是由鞘脂酶缺陷造成的，这种病的主要症状为肝大、脾大和严重弱智。病人一般在幼儿早期就死亡。

二、磷脂的合成

磷脂有甘油磷脂和鞘磷脂之分，它们在合成机制上不完全相同，但基本步骤相似，都包括 4 个阶段的反应：①骨架分子即甘油或神经鞘氨醇的合成；②疏水尾巴脂酰基从脂酰 -CoA 转移到骨架分子上，以酯键或酰胺键相连；③亲水性头部基团的加入，以磷酸酯键相连；④在某些情况下，头部基团发生修饰反应或者进行基团交换，以形成最后的磷脂分子。有时将这个阶段的反应称为磷脂合成的补救途径。

(一) 甘油磷脂的合成

Quiz4　两条合成途径的能耗是否相同?

甘油磷脂的合成是在光面内质网面向细胞质基质一侧的膜上进行的，其中第一阶段和第二阶段的反应与脂肪的合成完全一样(参看脂肪的合成)，而第三阶段的反应有两种方式。第一种方式：首先是磷脂酸被激活成 CDP- 二酰甘油(CDP-DG)，然后在相应的合酶催化下，与非活化的头部 X 基团起反应，生成各种甘油磷脂，同时释放出 CMP(图 28-5)；第二种方式：首先是磷脂酸被水解成 1,2- 二酰甘油，X 基团被活化成 CDP-X，然后在相应的合酶催化下，两者反应产生甘油磷脂，同时释放出 CMP(图 28-5)。一般细菌使用第一种方式，真核细胞中两种方式都存在，其中磷脂酰肌醇(PI)(图 28-6)和心磷脂的合成使用第一种方式，而磷脂酰胆碱(PC)、磷脂酰乙醇胺(PE)和磷脂酰丝氨酸(PS)使用第二种方式。

Quiz5　将 2 分子 FFA、1 分子甘油、1 分子丝氨酸通过第二条合成途径转变成 1 分子磷脂酰甘油，需要消耗多少分子的 ATP？

头部基团的活化反应类似于糖原合成中葡萄糖单位的活化(参看第二十七章“糖原代谢”)，第一步也是由激酶催化的消耗 ATP 的反应。但它与葡萄糖的活化有两点不同：一是后面并无变位酶催化的反应，二是由 CTP 而不是 UTP 参与最后一步反应。

能被直接激活的头部基团包括丝氨酸、乙醇胺和胆碱，它们最终分别成为 PS、PE 和 PC 中的 X 基团。

图28-5 甘油磷脂合成途径Ⅰ

图28-6 甘油磷脂合成途径Ⅱ

磷脂在光面内质网膜上合成的时候，会受到磷脂爬行酶（lipid scramblase）的催化，不断从膜的外层转移到内层，以维持两层的对称性生长。此外，内质网膜上的磷脂还可以通过小泡运输转移到其他膜系统。

PE 和 PC 除了可以通过上述方式由相应的 X 基团从头合成以外，还可以通过补救途径由 PS 衍生而来。例如，PS 经脱羧反应可转变为 PE，而 PE 从 *S*- 腺苷甲硫氨酸（SAM）接受甲基可转变为 PC。反过来，PS 可以通过 Ser 与 PE 上的乙醇胺通过基团交换的方式形成。

缩醛磷脂可视为一类特殊的甘油磷脂，与一般甘油磷脂的不同之处在于，与其甘油 1 号位羟基相连的基团并非脂酰基，而是通过醚键相连的、在 1 号位含有双键的烷基。由于含有特殊的醚键，因此它的合成与普通的甘油磷脂有所差别。

（二）鞘磷脂的合成

机体内的组织均可合成鞘磷脂，但以脑组织最为活跃，因为它是构成神经组织膜（特别是髓鞘）的主要成分，同样在光面内质网上合成。

e28-1 缩醛磷脂的合成

反应历程参看图 28-7，4 个阶段的反应是：①骨架分子二氢鞘氨醇（sphinganine）的合成，反应的原料是软脂酰 -CoA 和 Ser，还需要 NADPH；②疏水尾巴脂酰基的加入，形成 *N*- 脂酰基二氢鞘氨醇（*N*-acylsphinganine）；③*N*- 脂酰基二氢鞘氨醇经脱氢转变为 *N*- 脂酰基鞘氨醇（*N*-acylsphingosine）；④亲水头部基团的附着。

Quiz6 为什么膜脂的合成需要在膜上（主要是光面内质网）进行？

（三）古菌磷脂的合成

构成古菌的磷脂具有 3 个不同于细菌和真核生物的性质（参看第十六章“脂质与生物膜”）：①疏水的尾巴通过醚键而不是酯键与甘油骨架相连；②疏水尾巴衍生于异戊二烯单位而不是脂肪酸的碳氢链，因此具有分支；③甘油磷酸骨架的手性为 S 型，不同于细菌和真核生物的 R 型。有关古菌磷脂合成的具体步骤至今仍然不明。

$CH_3(CH_2)_{14}-CO-CoA$ + $H_3N^+-CH(COO^-)-CH_2OH$

软脂酰-CoA　丝氨酸

3-酮鞘氨醇合酶（磷酸吡哆醛）

CO_2

CoA

$CH_3(CH_2)_{14}-CO-CH(NH_3^+)-CH_2OH$

3-酮鞘氨醇

3-酮鞘氨醇还原酶

$NADPH+H^+$　$NADP^+$

$CH_3(CH_2)_{14}-CH(OH)-CH(NH_3^+)-CH_2OH$

二氨鞘氨醇

转酰酶

$CH_3(CH_2)_n-CO-CoA$

脂酰-CoA

CoA

$CH_3(CH_2)_{14}-CH(OH)-CH(NH-CO-(CH_2)_n-CH_3)-CH_2OH$

N-脂酰基二氢鞘氨醇

$FADH_2$　FAD

$CH_3(CH_2)_{12}-CH=CH-CH(OH)-CH(NH-CO-(CH_2)_n-CH_3)-CH_2OH$

胞苷

胆碱

$O=P(O^-)(O\text{-胞苷})-O-P(=O)(O^-)-OCH_2CH_2N^+(CH_3)_3$

CDP-胆碱

CMP

$CH_3(CH_2)_{12}-CH=CH-CH(OH)-CH(NH-CO-(CH_2)_n-CH_3)-CH_2O-P(=O)(O^-)-OCH_2CH_2N^+(CH_3)_3$

鞘磷脂

图 28-7　鞘磷脂的合成

第三节　糖脂代谢

一、糖脂的分解

鞘糖脂的水解发生在溶酶体，有一系列的酸性水解酶参与此过程。如图 28-8 所示，整个水解反应是高度有序的，当最后一步水解反应开始的时候，第一步反应即被关闭。如果有哪一种水解酶有缺陷，就会导致相应的遗传性疾病。例如，戈谢病（Gaucher disease）又称葡糖脑苷脂沉积病（glucocerbroside lipidcsis），是因 β- 葡糖脑苷脂酶（β-glucocerehrosidase）减少或缺失，使葡糖脑苷脂（glucocerbroside）不能分解成葡萄糖和 *N*- 酰基鞘氨醇，因而葡糖脑苷脂在各器官中大量沉积，引起组织细胞大量增殖。

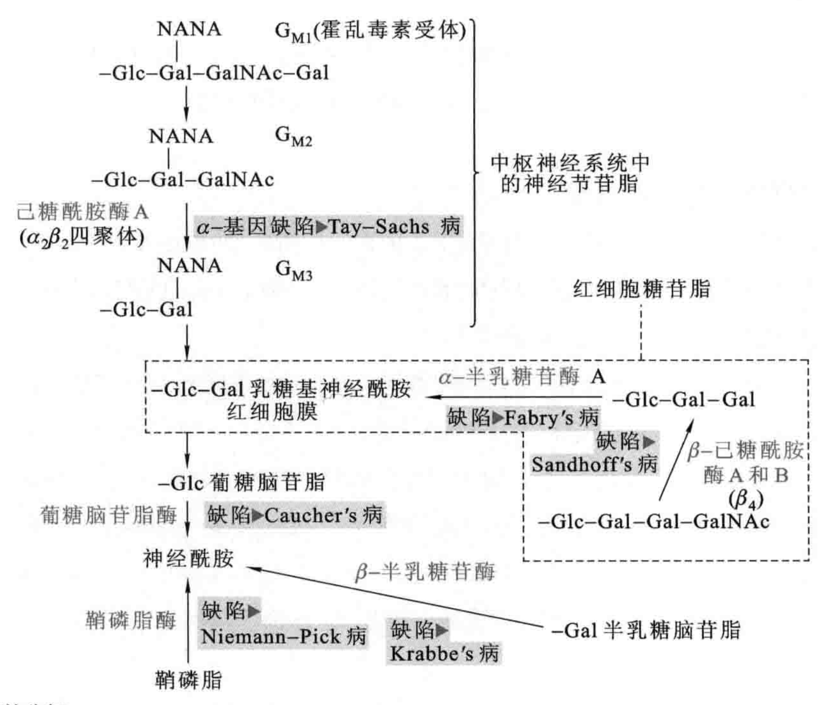


图 28-8　糖脂的分解

二、糖脂的合成

糖脂也分为甘油糖脂和鞘糖脂，现分别介绍它们的合成。

（一）鞘糖脂的合成

鞘糖脂合成的多数反应与鞘磷脂的合成一样，只是在神经酰胺形成以后，就发生一系列的糖基转移反应，催化转移反应的酶通称为糖基转移酶。当神经酰胺 1 号位的羟基接受一个葡糖残基以后，就形成最简单的鞘糖脂—葡糖脑苷脂。其他较为复杂的鞘糖脂（如各种神经节苷脂）是在此基础上由糖链进一步延伸而成的。

在糖基转移反应中，糖基供体为活化的单糖单位，通常为 UDP- 单糖。例如 UDP-Glc、UDP-Gal 和 UDP-GalNAc。少数为 CMP- 单糖，例如胞苷单磷酸 *N*- 乙酰神经氨酸（CMP-NeuAc）（图 28-9）。

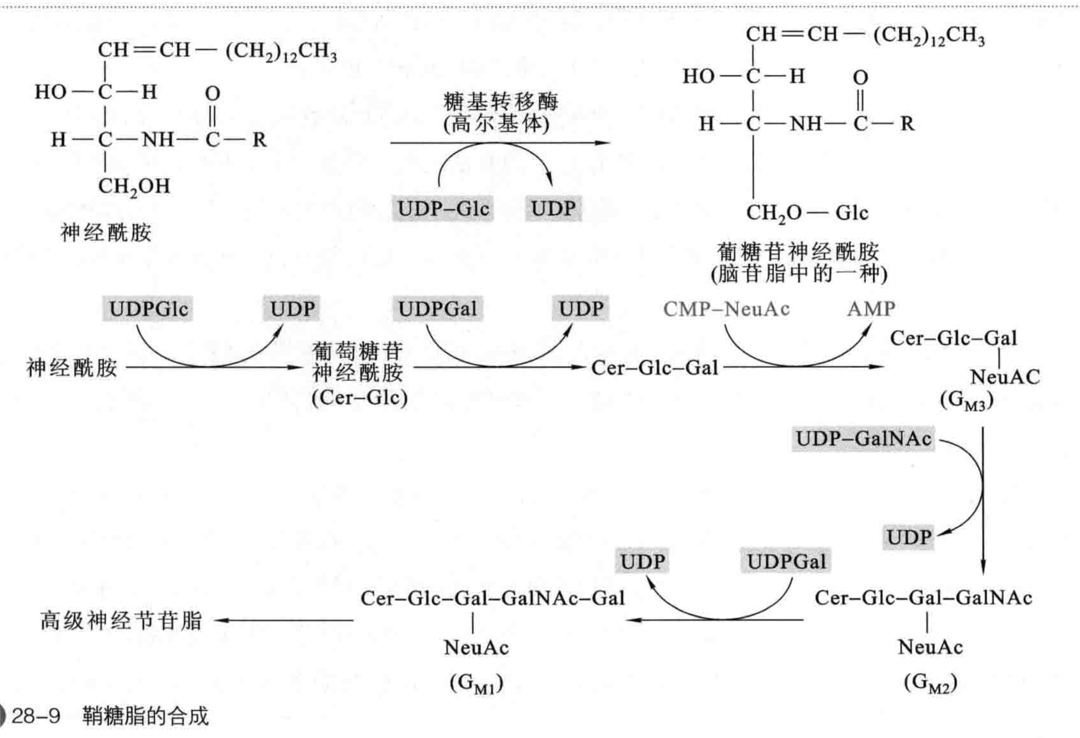


图 28-9　鞘糖脂的合成

（二）甘油糖脂的合成

甘油糖脂的合成与甘油磷脂合成的多数反应一样，只是在1，2- 二酰甘油形成以后，需要在特定的糖基转移酶催化下，甘油3号位的羟基从活化的单糖单位接受单糖基。

科学故事——脂瘦素及其受体的发现

肥胖是许多慢性病酝酿的温床，也是众多爱美之士的头号宿敌。20世纪90年代中期，科学家找到了两个与肥胖有关的基因，并证实它们具有减轻体重的功能。这个发现不仅是基础医学研究上一项重大突破，也可能在此基础上发展出一种彻底减肥的方法。

早在20世纪70年代，科学家就筛选出几株单基因突变造成的肥胖鼠，其中两株命名为*ob*(肥胖型，obese)及*db*(糖尿病型，diabetes)的肥胖鼠在所有肥胖特征上几乎完全一样，只有当与正常鼠"联体共生"(parabiosis)时，才会呈现不同的反应。所谓"联体共生"技术是利用手术将两只老鼠的循环系统结合在一起，使血液中的可溶性蛋白质因子相互流通。与正常鼠联体以后，*ob/ob*(表示两套染色体上的*ob*基因都突变)鼠的体重恢复正常，而*db/db*鼠却不会减轻体重。1973年，从事这项研究的Coleman猜测，*ob/ob*鼠由于基因突变，其血液中缺乏一种可调节体重的循环因子，而*db/db*鼠所突变的基因则可能负责接受该信号因子。这项假设在20年之后被证实。

1994年底，美国洛克菲勒大学的Freidman领导的研究小组，花了近8年的时间，利用定向克隆技术(positional cloning)，逐步寻找老鼠第六号染色体上的基因，终于找到了肥胖基因(*ob*)。该基因可翻译出长167个氨基酸、大小约1.6×10^4的蛋白质OB(后被命名为脂瘦素)。从结构上来看，它的N端具有一段典型的信号序列，表明脂瘦素是分泌蛋白。而在*ob*肥胖鼠体内的*ob*基因上，发现一个$C \rightarrow T$的突变，使得第105个氨基酸Arg变成终止氨基酸，而使肽链缩短无法表达出具有功能的产物。此外，*ob*基因的mRNA只有在脂肪组织中表达。从这些特征推测，脂瘦素是由脂肪细胞合成后进入循环系统，属于一种激素。它的功能可能是用来维持体内脂肪含量的恒定。也就是说，当体内脂肪贮量过高时，脂肪细胞便分泌脂瘦素进入血液，通报身体的"摄食控制中枢"(已有证据表明是位于下丘脑)，该中心便下达减少摄食及增加运动的指令，以消耗多余的脂肪。

1995年7月，Freidman、Amgen公司及Roche药厂三个不同的研究小组，在同一期*Science*上证实脂瘦素在老鼠体内具有调节体重的功能。不出所料，在正常老鼠血清中含有脂瘦素，但在*ob/ob*鼠血清中检测不到。将大肠杆菌表达出来的脂瘦素每天持续地注射到*ob/ob*鼠体内，二周后可使老鼠体重减轻约30%，若持续一个月可减轻达40%；同样的处理也可使正常鼠减重约12%。

进一步的研究还发现，老鼠体重减轻的质量相当于脂肪组织所减少的质量，表示减去的部分大都是老鼠体内的脂肪。这些经注射脂瘦素的老鼠相较于未注射的对照组，明显地摄食量减少，同时也更加活泼好动，显示出脂瘦素至少通过降低食欲及消耗能量两种方式，以达到减肥的目的。令人兴奋的是，将脂瘦素注射到一种摄取高脂肪食物诱导产生的肥胖鼠体内，居然也有减重的效果，这为开发脂瘦素成为减肥药，奠定了一个良好的基础。

然而，当Freidman的研究小组将脂瘦素打入*db/db*鼠体内以后，发现并没有减重的效果，同时他们还发现*db/db*鼠血清中的脂瘦素含量较正常鼠高。因此更加相信*db*基因的产物，可能就是脂瘦素的受体(OB-receptor，OB-R)。

与此同时，另一个位于美国马萨诸塞州的Millennium药厂研究小组也在如火如荼地寻找脂瘦素受体的基因。因为有报告指出大多数肥胖的人，血中脂瘦素的含量并不低，问题似乎出在它的受体基因上，或其下游信号转导的某个环节。在1995年底，脂瘦素的受体被找到了。他们利用放射性碘标记的脂瘦素为探针，在各种细胞株与组织切片中寻找可与之结合的蛋白质。结果发现在第三脑室及侧脑室的脉络丛有明显结合点(下丘脑也有，较不明显)，并从该组织成功地筛选出脂瘦素受体的cDNA。这个cDNA所翻译

出的蛋白质含有893个氨基酸，包括816个氨基酸残基的胞外结构域、23个氨基酸残基的跨膜区及34个氨基酸残基的胞内结构域。

从脂瘦素受体胞外结构域的结构上分析，可将其归为第一类细胞因子受体家族。令人疑惑的是，与其他同家族的细胞因子受体相比，脂瘦素受体的胞内结构域要小得多。如此小的结构域能正常地转导信号吗？另外，虽然脂瘦素受体cDNA被定位到老鼠第4号染色体上，与先前推测*db*基因的位置非常靠近，但从*db/db*鼠所筛选出的脂瘦素cDNA序列与正常的完全一样，并没有发现任何形式的突变，难道脂瘦素受体的基因并不是*db*基因？

为了抢先解开上述谜团，Millennium药厂、Freidman和美国哥伦比亚大学Rudolph L. Leibel等对此展开了激烈的竞争，3个月后终于真相大白。他们分别在*Cell*、*Nature*和*Science*上发表了他们各自的研究成果，都证实脂瘦素受体的基因就是*db*基因，并合理解释了在*db/db*鼠内脂瘦素受体的突变机制。原来Millennium药厂所找到的脂瘦素受体的cDNA，只不过是同一种mRNA前体的6种不同可变剪接型的一种(参看第四十三章"真核生物的基因表达调控")，即OB-Ra。另一种剪接类型存在于下丘脑，是已在人类细胞中发现的OB-Rb型，其翻译产物具有302个氨基酸的胞内结构域，可能才是有功能的脂瘦素受体。

从*db/db*鼠所得到的OB-Rb cDNA多出了一段106 nt的插入序列，这段序列中含有一个终止密码子，迫使翻译提前结束，只能表达出具有较短胞内区的OB-R蛋白。而造成106 nt插入的原因是染色体DNA上一个G→T的点突变，产生一个剪接供体位点(参看第三十七章"转录后加工")，使得位于突变点前的序列被保留到OB-Rb的mRNA。

e28-2 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/lipid-synthesis.html#phospholipids(完全免费的医学生物化学课程网站，有关磷脂代谢的内容)
2. http://www.wormbook.org/chapters/www_obesity/obesity.html(书虫网站有关脂代谢与肥胖的内容)
3. http://www.med.uiuc.edu/m1/biochemistry/TA%20reviews/phospholipids.htm(美国伊利诺伊大学厄本那－香槟分校有关磷脂代谢的内容，以问答的形式展开)

参考文献：

1. Münzberg H, et al. Mechanisms of leptin action and leptin resistance. *Annu Rev Physiol*, 70: 537–556, 2008.
2. Friedman J M, Halaas J L. Leptin and the regulation of body weight in mammals. *Nature*, 395(6704): 763–770, 1998.
3. Raetz C R H, Dowhan W. Biosynthesis and function of phospholipids in *Escherichia coli*. *J Biol Chem*, 265: 1 235–1 238, 1990.
4. Bishop W R, Bell R M. Assembly of phospholipids into cellular membranes: biosynthesis, transmembrane movement and intracellular translocation. *Annu Rev Cell Biol*, 4: 579–610, 1988.

第二十九章　脂肪酸代谢

作为多种脂质的组分，脂肪酸在细胞中具有多项功能，但最重要的功能是为生物膜的主要成分磷脂提供疏水尾巴，并且作为储能物质与甘油缩合成脂肪，贮存在脂肪组织中，当机体需要的时候进行氧化分解，为机体供能。机体在饥饿的情况下，脂肪酸氧化分解加速，分解产生的乙酰 -CoA 还可以在肝细胞的线粒体中，转化为小分子的酮体，而为脑、骨骼肌和心肌提供一种替代的能源；相反，机体在能量状态较高的时候，脂肪酸合成会加快，以便让富余的能量能及时得到储存。

本章将主要介绍脂肪酸的氧化分解、合成以及相关的调节机制。

第一节　脂肪酸的分解

脂肪酸的分解是以氧化的形式进行的，而氧化方式又分为 α- 氧化、β- 氧化和 ω- 氧化，其中 β- 氧化是主要的方式。现分别加以讨论。

一、脂肪酸的 β- 氧化

（一）β- 氧化的发现

对脂肪酸降解的研究有一段漫长而有趣的历史，早在 20 世纪初，德国科学家 Franz Knoop 开创性地使用难以被动物代谢的苯环标记脂肪酸，并将标记的脂肪酸分成偶数脂肪酸和奇数脂肪酸两组，然后分别喂养狗。经过一定时间以后，从狗尿中分析脂肪酸的降解产物。结果发现：凡是偶数脂肪酸的最终降解产物均为苯乙酸（phenylacetate）的衍生物，而奇数脂肪酸的最终降解产物都是苯甲酸（benzoate）的衍生物（图 29-1）。Knoop 由此得出，脂肪酸的降解是分步进行的，每一步去除一个二碳单位，而起始的氧化反应发生在 β- 碳原子上，这就是他提出的脂肪酸 β- 氧化学说。

图 29-1　Knoop 的苯环标记实验

Knoop 的实验具有特殊的意义，这是人类第一次使用示踪的方法研究物质代谢。如果今天让我们重复他的实验，肯定会首选同位素去标记脂肪酸，但在那个年代同位素还没有被人类发现。

后来，Albert Lehninger 证明了 β- 氧化发生在线粒体基质，而 F. Lynen 和 E. Reichart 则进一步确定了释放出来的"二碳单位"为乙酰 -CoA，而不是游离的乙酸。

（二）β- 氧化的反应历程

1. 脂肪酸的活化

正如葡萄糖在氧化之前需要被活化成 6- 磷酸葡糖一样，脂肪酸在进行 β- 氧化之前也需要被活化。脂肪酸的活化形式为脂酰 -CoA，催化活化反应的酶是脂酰 -CoA 合成酶（acyl-CoA synthetase），此

酶也被称为硫激酶(thiokinases)。

脂酰 -CoA 合成酶催化的反应机制参看图 29-2:首先是脂肪酸与 ATP 形成脂酰 -AMP 中间物,同时释放出 PPi。随后 CoA 置换出 AMP,形成脂酰 -CoA,而先前释放出来的焦磷酸被焦磷酸酶水解成无机磷酸,使得总反应的 $\Delta G^{\ominus\prime}$ 为一个大的负值,脂肪酸的活化反应也因此是不可逆的。在这一步反应中,ATP 转变成 AMP 和 PPi,而 PPi 又迅速被焦磷酸酶水解,因此,一般认为每活化 1 分子游离的脂肪酸需要消耗 2 分子 ATP。

图 29-2 脂酰 -CoA 合成酶的催化机理

Quiz1 试设计一个实验,证明脂肪酸的活化是两步反应。

有几类不同的脂酰 -CoA 合成酶,它们对脂肪酸碳链的长度具有不同的要求。其中有一类位于线粒体的外膜,专门在细胞质基质激活长链脂肪酸(12 碳 ~18 碳);还有一类位于线粒体内膜,负责激活直接从细胞质进入线粒体基质的中、短链脂肪酸(2 碳 ~10 碳)。

2. 脂酰 -CoA 的转运

由于长链的脂酰 -CoA 不能直接透过线粒体内膜,因此需要专门的转运系统将其转变为基质内的脂酰 -CoA。

构成脂酰 -CoA 跨膜转运系统的组分包括肉碱 : 软脂酰转移酶Ⅰ(carnitine palmitoyltransferase Ⅰ, CPT- Ⅰ)、转位酶(translocase)和肉碱 : 软脂酰转移酶Ⅱ(CPT- Ⅱ)。

肉碱的化学名称为 L-β- 羟基 -γ- 三甲铵丁酸,它在转运系统中作为脂酰基的中间受体起作用。

如图 29-3 所示,CPT- Ⅰ位于线粒体外膜,当细胞质基质的脂酰 -CoA 进入膜间隙以后,在 CPT- Ⅰ的催化下,脂酰基被转移到肉碱的 β- 羟基上,形成脂酰 - 肉碱,CoA 则被游离出来并返回到细胞质基质。随后,脂酰 - 肉碱在线粒体内膜特定的转位酶帮助下进入基质。在位于内膜基质一侧的 CPT- Ⅱ催化下,脂酰基被转移到基质 CoA 的巯基上重新生成脂酰 -CoA,而肉碱又经转位酶反向运输出基质。

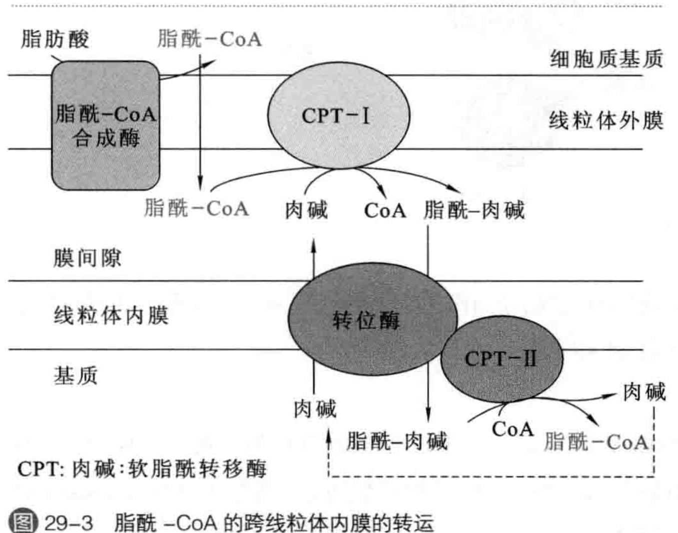


图 29-3 脂酰 -CoA 的跨线粒体内膜的转运

由此可见,肉碱在脂肪酸氧化降解过程中所起的作用是不可低估的。如果肉碱不足,就会影响到细胞内脂肪酸的氧化,致使骨骼肌、心肌和肝这些需要脂肪酸氧化供能的组织发生病变。正常的人一般很少缺乏肉碱,因为一方面可以从膳食中摄取肉碱,以肉类和乳制品最丰富,另一方面自己也可以合成。合成部位主要是肝细胞,需要维生素 C、B_6 和 PP 以及铁作为辅助因子,前体包括 Lys 和 Met。因此,将肉碱视为减肥药物是缺乏科学依据的。

Quiz2 对先天缺乏肉碱的人,你对他们摄入脂肪酸的种类有何好的建议?

当细胞质基质中的脂酰-CoA转化为线粒体基质的脂酰-CoA以后，脂肪酸的β-氧化便可真正地开始了。此后将循环发生脱氢、加水、再脱氢和硫解(thiolysis)反应，直到脂酰-CoA的碳链完全氧化裂解成乙酰-CoA(图29-4)。反应中形成的NADH和$FADH_2$进入呼吸链，而乙酰-CoA进入TCA循环进一步氧化分解。

Quiz3 脂肪酸β-氧化的前三步反应与TCA循环中的哪三步反应相似?

$RCH_2CH_2—C(=O)—S—CoA$ 脂酰-CoA

脂酰-CoA脱氢酶 (FAD → $FADH_2$)

$RCH=CH—C(=O)—S—CoA$ 反烯酰-CoA

烯酰-CoA水合酶 (H_2O)

$RCH(OH)—CH_2—C(=O)—S—CoA$ L-β-羟酰-CoA

羟酰-CoA脱氢酶 (NAD^+ → NADH + H^+)

$RC(=O)—CH_2—C(=O)—S—CoA$ β-酮酰-CoA

硫解酶 (CoASH)

$RC(=O)—S—CoA$ 脂酰-CoA+乙酰-CoA $CH_3—C(=O)—S—CoA$

图29-4 一轮β-氧化循环的四步反应

3. 脱氢

这是一步由脂酰-CoA脱氢酶(acyl-CoA dehydrogenase)催化的氧化还原反应，FAD为电子受体。从一种叫阿开木(ackee)的植物中提取的降糖氨酸(hypoglycin)在体内的代谢物是该酶的强抑制剂。脱氢反应是高度立体专一性的，产物是Δ^2-反烯酰-CoA(*trans*-Δ^2-enoyl-CoA)和$FADH_2$。需要特别注意的是，脂肪酸β-氧化产生的$FADH_2$并非通过复合体Ⅱ，而是依次通过电子传递黄素蛋白(electron-transferring flavin, ETF)、铁硫蛋白和CoQ进入呼吸链(图29-5)，然后直通复合物Ⅲ。因此1分子$FADH_2$产生1.5分子ATP。

Quiz4 降糖氨酸具有降血糖的作用，因此有人用它来治疗糖尿病，但也容易引起中毒，可导致服用它的人不能忍受长时间的饥饿。对此你如何解释?

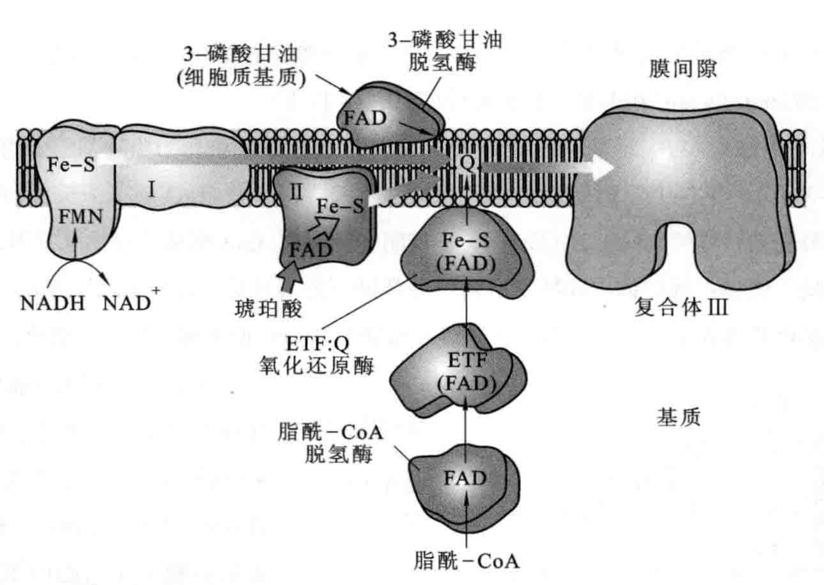


图29-5 脂酰-CoA脱氢酶与呼吸链之间的联系

已经发现3种不同类型的脂酰-CoA脱氢酶，它们分别作用于长链脂酰-CoA(12碳~18碳)、中链脂酰-CoA(8碳~12碳)和短链脂酰-CoA(4碳~8碳)。

4. 加水

这是一步由烯酰-CoA水合酶(enoyl-CoA hydratase)催化的反应，H_2O作为底物参加反应。反应也是高度立体专一性的，被水合的双键只能是反式，而生成的产物只会是L-羟酰-CoA(L-hydroxyacyl-coenzyme A)，并且羟基一定是加在β-碳原子上。

5. 再脱氢

这又是一步氧化还原反应，由羟酰 -CoA 脱氢酶（hydroxyacyl-CoA dehydrogenase）催化，被氧化的是 β- 碳原子，NAD^+ 为电子受体，产物为 β- 酮酰 -CoA（β-ketoacyl-CoA）和 NADH。后者直接从复合体 I 进入呼吸链，1 对电子产生 2.5 分子 ATP。

6. 硫解

这是一步由硫解酶（thiolase）催化的反应，反应机制如下：酶活性中心的一个 Cys 残基用它的巯基亲核进攻 β- 羰基碳，释放出乙酰 -CoA，同时形成少了 2 个碳原子的脂酰 - 酶中间物。随后，CoA 上的巯基取代酶分子 Cys 上的巯基，产生新的脂酰 -CoA（少 2 个碳）。

（三）β- 氧化小结

以 1 分子软脂酸为例，需要经过 7 轮 β- 氧化循环，共产生 8 分子乙酰 -CoA、7 分子 $FADH_2$ 和 NADH，总反应式为：

$$\text{软脂酰 -CoA} + 7FAD + 7NAD^+ + 7H_2O \longrightarrow 8\ \text{乙酰 -CoA} + 7FADH_2 + 7NADH + 7H^+$$

其完全氧化可以产生 106 分子 ATP（表 29-1）。

表 29-1　1 分子软脂酸彻底氧化以后 ATP 的收支情况

与 ATP 产生有关的酶	NADH 或 $FADH_2$ 产生的量	最终产生 ATP 的数目
脂酰 -CoA 合成酶		–2
脂酰 -CoA 脱氢酶	7 $FADH_2$	7 × 1.5=10.5
羟酰 -CoA 脱氢酶	7 NADH	7 × 2.5=17.5
异柠檬酸脱氢酶	8 NADH	8 × 2.5=20
α- 酮戊二酸脱氢酶	8 NADH	8 × 2.5=20
琥珀酰 -CoA 合成酶		8 GTP=8 ATP
琥珀酸脱氢酶	8 $FADH_2$	8 × 1.5=12
苹果酸脱氢酶	8 NADH	8 × 2.5=20
总量		106

（四）β- 氧化的生理功能

脂肪酸 β- 氧化的主要生理功能是产生 ATP，其效率要高于葡萄糖。例如，含有 12 个碳原子的 1 分子饱和脂肪酸完全氧化，可以产生 78 分子 ATP，而 2 分子的葡萄糖（共 12 个碳原子）完全氧化只能产生 60~64 分子的 ATP。

Quiz5　1 分子乙酸或乙醇在人体内完全氧化可分别产生多少分子的 ATP？

脂肪酸 β- 氧化的另外一个重要功能是与呼吸链一起为机体产生代谢水（metabolic water）。虽然 β- 氧化在水合反应中消耗了水分子，但产生的乙酰 -CoA、$FADH_2$ 和 NADH 经彻底氧化可以产生更多的水分子。这对于某些生活在干燥缺水环境的生物十分重要，像骆驼和长鼻袋鼠（kangaroo rat）将 β- 氧化作为它们获取水源的一种特殊手段。

（五）奇数脂肪酸的 β- 氧化

奇数脂肪酸虽然在自然界的含量远远低于偶数脂肪酸，但它也存在于生物体，在某些生物（如反刍动物）体内，其氧化放出的能量可占它们所需能量的 25%。此外，它还可以作为动物糖异生的前体。因此，奇数脂肪酸在能量代谢中具有特别的作用。

奇数脂肪酸的氧化实际上就是丙酰 -CoA 的氧化，因为碳原子数目 ≥5 的奇数脂肪酸完全可以和偶数脂肪酸一样进行 β- 氧化，直到丙酰 -CoA（propionyl-CoA）出现为止。

丙酰 -CoA 的氧化原则上并不困难，只需将它转变为偶数的琥珀酰 -CoA 即可。如图 29-6 所示，丙酰 -CoA 先后在丙酰 -CoA 羧化酶（propionyl-CoA carboxylase）、甲基丙二酸单酰 -CoA 消旋酶（methyl-malony-CoA racemase）和甲基丙二酸单酰 -CoA 变位酶（methyl-malony-CoA mutase）的催化下，最终转变为琥珀酰 -CoA，而琥珀酰 -CoA 可顺利进入 TCA 循环被进一步氧化分解，或者转变为草酰乙酸以后离开循环而作为糖异生前体。

Quiz6　戊酸和丁酸在人体内彻底氧化，哪一种产生的 ATP 更多？

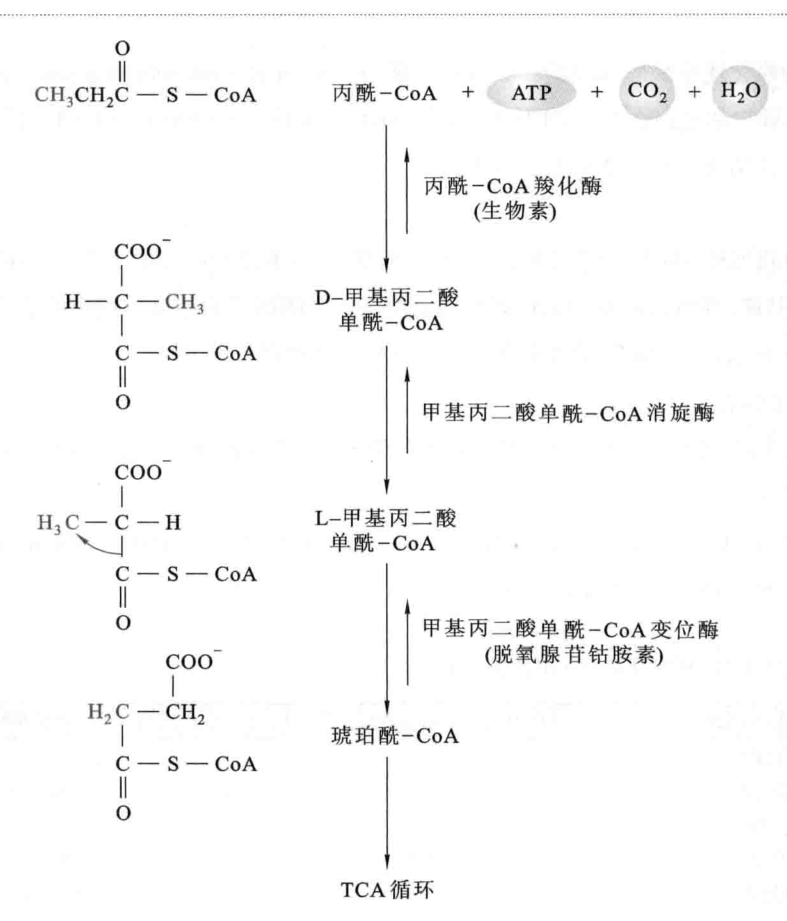


图 29-6 丙酰 -CoA 的氧化利用

Quiz7 为什么严格的素食者利用奇数脂肪酸的能力减弱?

值得注意的是,甲基丙二酸单酰 -CoA 变位酶需要脱氧腺苷钴胺素作为辅酶,而这种辅酶衍生于维生素 B_{12},所以,维生素 B_{12} 的缺乏会影响到机体对奇数脂肪酸的代谢。

(六) 不饱和脂肪酸的 β- 氧化

不饱和脂肪酸在进行 β- 氧化的时候会遇到特别的难题,即如何处理它本身位置不对的顺式双键。但一般说来,这些双键本来离羧基端有一段距离,因此一开始照常进行 β- 氧化,只是当顺式双键进入 β 位以后,β- 氧化将不能继续进行。这时需要特殊的异构酶即烯酰 -CoA 异构酶(enoyl-CoA isomerase)来改变双键的位置和性质,使之转变为可被脂酰 -CoA 脱氢酶识别的 2 号位的反式双键,β- 氧化就得以继续了(图 29-7)。

e29-1 多不饱和脂肪酸的 β- 氧化

上述烯酰 -CoA 异构酶只能解决单不饱和脂肪酸在进行 β- 氧化时遇到的问题,如果是多不饱和脂肪酸,还需要 2,4- 二烯酰 -CoA 还原酶(2,4-dienoyl-CoA reductase)的帮助。

需要特别指出的是,大于 18 个碳原子的脂酰 -CoA 难以进入线粒体进行 β- 氧化,但可以进入过氧化物酶体或乙醛酸循环体进行 β- 氧化。脂酰 -CoA 若要进入过氧化物酶体或乙醛酸循环体,需要膜上的一种转运蛋白,但不需要肉碱。

e29-2 超长链脂肪酸的 β- 氧化

二、脂肪酸的 α- 氧化与 ω- 氧化

脂肪酸 α- 氧化是指发生在 α- 碳原子上的氧化。像植烷酸(phytanic acid)这样的脂肪酸,因在 β- 碳原子上有甲基,故在细胞内难以直接进行 β- 氧化,必须先通过 α- 氧化去除 1 个碳原子。

Quiz8 为什么脂肪酸 α- 氧化不能在脂酰 -CoA 上进行?

与 β- 氧化不同的是,α- 氧化直接在游离的脂肪酸上进行,不需活化,不需消耗 ATP,但也不产生 ATP。整个 α- 氧化既可以发生在内质网,也可以发生在线粒体或过氧化物酶体。下面以植烷酸为例,对 α- 氧化做简单的介绍。

植烷酸来源于反刍动物脂肪和牛奶,叶绿素的组分叶绿醇在体内也可以转变为植烷酸。如

$CH_3(CH_2)_7-CH=CH-CH_2(CH_2)_6C(=O)-S-CoA$ 油酰-CoA

↓ 三轮β-氧化循环 → $3CH_3-C(=O)-CoA$

$CH_3(CH_2)_7CH=CH-CH_2-C(=O)-S-CoA$ （γ，β） Δ^3-顺-十二烯酰-CoA

⇅ 烯酰-CoA异构酶

$CH_3(CH_2)_7CH_2-CH=CH-C(=O)-S-CoA$ （β，α） Δ^3-反-十二烯酰-CoA

↓ H_2O 烯酰-CoA水合酶

$CH_3(CH_2)_7CH_2-CH(OH)-CH_2-C(=O)-S-CoA$

↓ β-氧化重新进行

$6\,CH_3C(=O)-S-CoA$

图 29-7　单不饱和脂肪酸（油酸）的β-氧化

图 29-8 所示，首先在α-羟化酶（α-hydroxylase）催化下，植烷酸的α-碳原子发生羟基化反应。这里的α-羟化酶需要O_2、Fe^{2+}和维生素 C 的参与。随后，在α-氧化酶的催化下，α-碳原子被氧化成羰基，而原来的羧基变成了CO_2。经过以上两步反应，植烷酸少了一个碳原子，变成了降植烷酸（pristanic acid），原来位于β-碳原子上的甲基"移花接木"到了α-碳原子上。于是，降植烷酸就可以照常进行β-氧化了。

Quiz9　为什么植烷酸在人体内可以作为糖异生的前体?

由于人的正常饮食中含有大量植烷酸或其前体植烷醇，因此α-氧化系统的正常运行对于这类脂肪酸的代谢是必不可少的。有一种遗传病叫雷夫苏姆病（Refsum disease），患者先天缺乏α-氧化相关的酶，导致摄入的植烷酸在体内堆积，引起中毒，其主要症状包括视网膜色素变性（retinitis pigmentosa）、夜视能力减弱、外周神经疾病（peripheral neuropathy）和小脑性运动失调（cerebellar ataxia）等。目前这种疾病没有特别的治疗方法，只能通过严格控制食物中植烷酸的摄入来防止其积累。

脂肪酸除了可以进行α-氧化和β-氧化以外，还可以进行ω-氧化。脂肪酸的ω-氧化发生在末端甲基即ω-碳原子上，被氧化的脂肪酸与α-氧化一样不需要活化。

e29-3　脂肪酸的ω-氧化

三、酮体的生成和利用

在饥饿、禁食或某些病理状态下（如糖尿病）下，人体内的脂肪动员加强，大量的脂肪酸被肝细胞吸收和氧化。与此同时，为了维持血糖浓度的稳定，体内的糖异生作用也被激活。草酰乙酸作为糖异生的原料被消耗，导致肝细胞内草酰乙酸浓度急剧下降，进而影响到三羧酸循环，大量由脂肪酸β-氧化产生的乙酰-CoA 因得不到及时氧化而出现堆积，酮体（ketone bodies）就在这种情况下生成了。

酮体包括丙酮、乙酰乙酸（acetoacetate）和 D-β-羟丁酸（β-hydroxybutyrate），其合成场所是肝细胞的线粒体基质。参与酮体合成的酶有硫解酶、β-羟-β-甲基戊二酸单酰-CoA 合酶或 HMG-CoA 合酶

(β-hydroxy-β-methylglutaryl CoA synthase)、HMG-CoA 裂合酶(HMG-CoA lyase) 和 D-β- 羟丁酸脱氢酶(D-β-hydroxybutyrate dehydrogenase)(图 29-9)。反应的前两步与胆固醇合成的第一步和第二步相同,只不过胆固醇合成的场所是细胞质基质,并且不限于肝细胞。

酮体合成的第一步反应为硫解酶的逆反应,第三步反应产生第一种酮体——乙酰乙酸,乙酰乙酸既可以自发脱羧基变成丙酮,又可以在脱氢酶的催化下被还原为 D-β- 羟丁酸。

Quiz10 人体细胞内为什么没有一个专门的酶催化乙酰乙酸的脱羧反应?

酮体产生以后,就可以通过自由扩散从肝细胞进入血液,随着血液循环到达肝外组织(如脑、骨骼肌、心肌和肺)。其中丙酮主要经肺呼出体外,乙酰乙酸和β- 羟丁酸在重新转变为乙酰 -CoA(图 29-10)后,可进入三羧酸循环氧化放能。因此,在饥饿或禁食的时候,酮体成为脑、骨骼肌和心肌一种极为重要的替代燃料。但在肝细胞内,酮体不能再转变为乙酰 -CoA,原因是肝细胞缺乏利用酮体的酶——β- 酮酰 -CoA 转移酶或乙酰乙酸硫激酶。

β- 酮酰 -CoA 转移酶也称为琥珀酰 -CoA 转硫酶,它催化琥珀酰 -CoA 上

图 29-8 植烷酸的氧化

图 29-9 酮体的形成

$$CH_3-\underset{H}{\overset{OH}{C}}-CH_2-COO^- \xrightleftharpoons[\text{D-}\beta\text{-羟丁酸脱氢酶}]{NAD^+ \rightarrow NADH+H^+} CH_3-\overset{O}{\overset{\|}{C}}-CH_2-COO^- \xrightarrow[\beta\text{-酮酰-CoA 转移酶}]{\text{琥珀酰-CoA} \rightarrow \text{琥珀酸}} CH_3-\overset{O}{\overset{\|}{C}}-CH_2-\overset{O}{\overset{\|}{C}}-SCoA \xrightleftharpoons[\text{硫解酶}]{CoASH} 2\,CH_3-\overset{O}{\overset{\|}{C}}-SCoA$$

D-β-羟丁酸　　乙酰乙酸　　乙酰乙酰-CoA

图 29-10　酮体的利用

的 CoA 转移到乙酰乙酸，致使乙酰乙酰 -CoA 重新形成。乙酰乙酸硫激酶催化乙酰乙酸与 ATP、CoA 反应，直接形成乙酰乙酰 -CoA，反应机制类似于脂酰 -CoA 合成酶。重新形成的乙酰乙酰 -CoA 在硫解酶的催化下，转变为乙酰 -CoA 并进入三羧酸循环。

乙酰乙酸和 β- 羟丁酸均为中强酸，因此酮体产生过多会引起酮症（ketosis），甚至出现酸中毒。

第二节　脂肪酸的合成

在动物体内偶数脂肪酸不能完全转变成糖，但是，糖却很容易变为偶数脂肪酸，并进一步转化为脂肪。生物体在能量充裕的情况下，会将一部分多余的能量以糖原的形式贮存起来。然而，糖原是一种短期能源储备，且动物体合成糖原的量是受到限制的，因此，当糖原积累达到极限的时候，机体只能依赖于脂肪了。脂肪作为一种长期能量储备，不但效率高，而且容易堆积在一起而不改变细胞的渗透压。可以说，生物体合成和积累脂肪的能力几乎是无限的。

脂肪酸合成的研究开始于 Knoop 提出 β- 氧化学说以后不久。最初，人们推测，脂肪酸合成可能是 β- 氧化的逆反应，但事实并非如此。就像糖异生和糖酵解之间的关系一样，脂肪酸合成采取的是一条与 β- 氧化不同的途径，两者在反应的场所、酶、电子受体或供体、脂酰基的载体或二碳单位的形式等方面都具有显著的差别（图 29-11）。

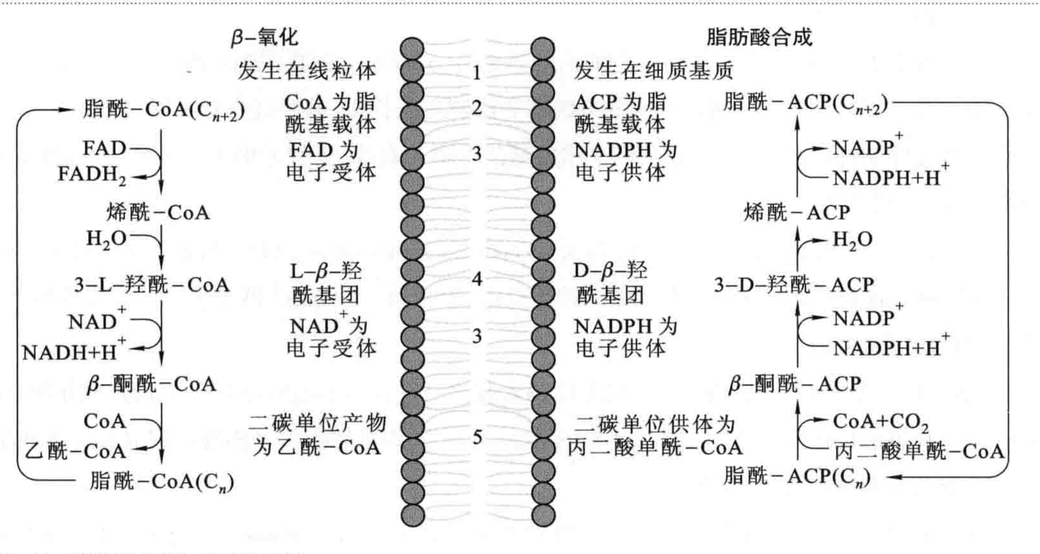


图 29-11　脂肪酸分解与合成的比较

一、脂肪酸合成的一般性质

几乎所有的生物都能够进行脂肪酸的合成，虽然在具体的细节上会有所差别，但至少有以下共同的性质：①除了植物是在质体内，其他生物合成的场所均为细胞质基质；②从头合成需要乙酰 -CoA 作为引物，若要合成奇数脂肪酸，就以丙酰 -CoA 作为引物；③丙二酸单酰 -CoA（malonyl-CoA）作为活化

的二碳单位供体；④丙二酸单酰 -CoA 的脱羧反应和 NADPH 作为驱动碳链延伸的动力；⑤软脂酸通常是反应的终产物；⑥软脂酸以外的脂肪酸是通过修饰形成的。

二、脂肪酸合成的详细机制

（一）乙酰 -CoA 跨线粒体内膜的转运

同位素示踪实验证明了脂肪酸合成的前体为乙酰 -CoA。对于真核生物而言，其体内形成乙酰 -CoA 的反应，不论是丙酮酸的氧化脱羧，还是脂肪酸的氧化或氨基酸碳骨架的转变，都主要发生在线粒体基质，而乙酰 -CoA 并不能直接透过线粒体内膜，但内膜上有一种柠檬酸 - 丙酮酸穿梭系统，可将基质内的乙酰 -CoA 转化成细胞质基质中的乙酰 -CoA，该系统作用的机制（图 29–12）是：基质内的乙酰 -CoA 先与草酰乙酸缩合成柠檬酸，然后，柠檬酸通过内膜上特殊的转运蛋白进入细胞质基质。一旦柠檬酸进入细胞质基质，其中的柠檬酸裂合酶（citrate lyase）就将其裂解成乙酰 -CoA 和草酰乙酸。草酰乙酸可以经过苹果酸和丙酮酸中间物重新转变为基质内的草酰乙酸。

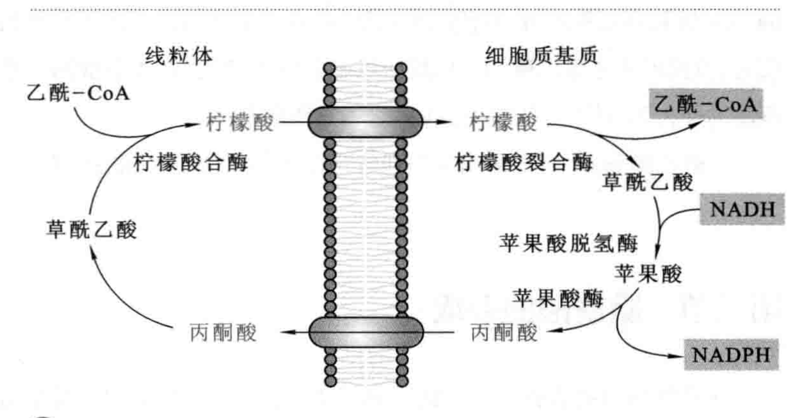


图 29–12 柠檬酸 - 丙酮酸跨膜穿梭系统

然而，柠檬酸作为三羧酸循环的中间物，在什么情况下会离开循环为脂肪酸合成提供原料呢？实际上这完全取决于细胞能量状态的高低。当细胞的能量状态很高时，ATP 作为负别构效应物抑制异柠檬酸脱氢酶的活性（参看第二十三章“三羧酸循环”），从而导致异柠檬酸的积累，异柠檬酸在顺乌头酸酶的催化下很容易转变为柠檬酸，这就导致了柠檬酸的积累。柠檬酸的积累为它离开线粒体进入细胞质基质创造了条件。

（二）乙酰 -CoA 的活化

人们在确定了乙酰 -CoA 是脂肪酸合成的前体之后，并没有立刻意识到乙酰 -CoA 需要先活化才能被利用。有人曾经长时间利用肝细胞抽取物研究脂肪酸的合成，但一直没有进展，直到有一天偶然用碳酸氢盐缓冲系统进行研究，才观测到了脂肪酸的合成。自此，人们才确定了乙酰 -CoA 的活化在脂肪酸合成中的重要性。

乙酰 -CoA 的活化是在乙酰 -CoA 羧化酶（acetyl–CoA carboxylase，ACC）的催化下完成的（图 29–13），反应机理类似于糖异生途径中丙酮酸的羧化反应（参看第二十五章“糖异生”）。该反应为不可逆反应，是脂肪酸合成的限速反应。

细菌的 ACC 由 3 个亚基组成，分别具有生物素载体（biotin carboxyl carrier）、生物素羧化酶（biotin carboxylase）和羧基转移酶的功能。真核生物的 ACC 是一个多功能酶，一条多肽链同时具有 3 种功能。

（三）脂肪酸合酶的结构与功能

参与脂肪酸合成的蛋白质或酶通称为脂肪酸合酶（fatty acid synthase），它包括：①酰基载体蛋白（acyl carrier protein，ACP）；②乙酰 -CoA：ACP 转脂酰基酶（acetyl-CoA-ACP transacylase，AT）；③丙二酸单酰 -CoA：ACP 转酰酶（malonyl-CoA-ACP transacylase，MT）；④ β- 酮酰 -ACP 合酶（β-ketoacyl-ACP synthase，KS），也称为缩合酶（condensing enzyme，CE）；⑤ β- 酮酰 -ACP 还原酶（β-ketoacyl-ACP reductase，KR）；⑥ β- 羟 -ACP 脱水酶（β-hydroxy-ACP dehydrase，DH）；⑦烯酰 -ACP 还原酶（enoyl-ACP reductase，ER）；⑧硫酯酶（thioesterase，TE）。

细菌和植物体内的脂肪酸合酶以多酶复合体形式存在，不同的酶活性由不同的蛋白质承担；而酵

乙酰-CoA碳负离子 + 羧化生物素 —乙酰-CoA羧化酶→ 丙二酸单酰-CoA + 生物素酸

BCCP　BCCP

H^+

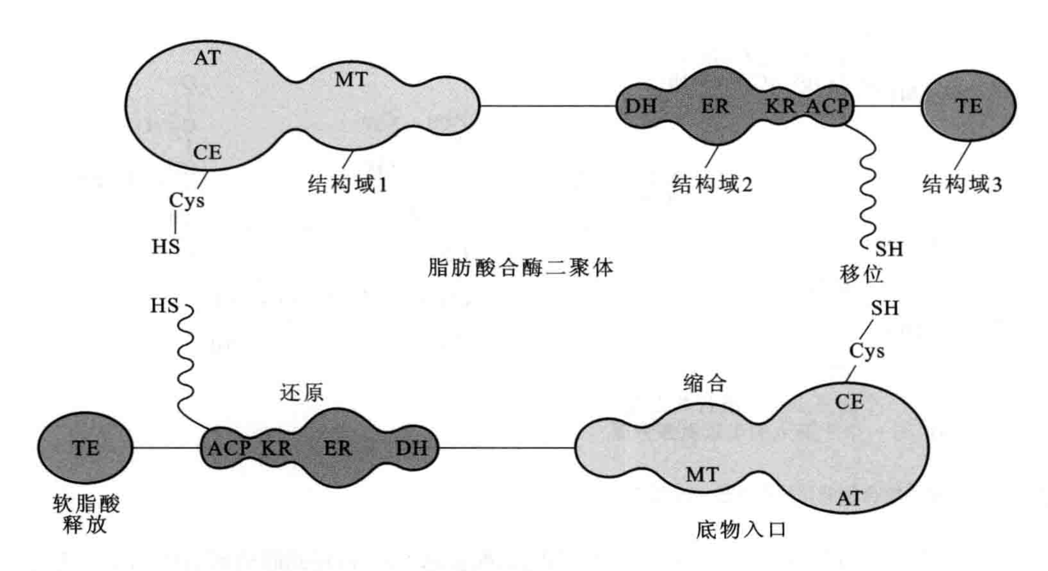
生物素

HCO_3^- + ATP —乙酰-CoA羧化酶→ ADP + Pi

图 29-13　乙酰 -CoA 的羧化反应

母细胞的脂肪酸合酶有 6 个 α 亚基和 6 个 β 亚基($\alpha_6\beta_6$)组成，其中 α 亚基具有 ACP、KS 和 KR 的活性，β 亚基具有 AT、MT、DH、ER 和 TE 的活性；哺乳动物的脂肪酸合酶则是一种多功能酶(图 29-14)，由两个相同的亚基头尾相连。每一个亚基有 3 个结构域，同时具有 ACP 的功能和 7 种酶的活性。但单个亚基并不能完成一个脂肪酸分子的合成。

Quiz11　为什么哺乳动物脂肪酸合酶的单个亚基不能完成一个脂肪酸的合成？



结构域1：AT=转酰酶；MT=丙二酸单酰转移酶；CE=缩合酶
结构域2：ACP=酰基载体蛋白；KR=β-酮酰COA还原酶；DH=脱水酶；ER=烯酰还原酶
结构域3：TE=硫酯酶

图 29-14　哺乳动物脂肪酸合酶的结构模型

ACP 在脂肪酸合成中的功能是作为脂酰基的载体，其功能端的结构与乙酰 -CoA 十分相似(图 29-15)，由磷酸泛酰巯基乙胺与它的一个 Ser 残基的羟基以磷酸酯键相连，因此，ACP 可视为在功能上放大的 CoA。

4- 磷酸泛酰乙胺

HS—CH_2—CH_2—NH—C(=O)—CH_2—CH_2—NH—C(=O)—CH(OH)—C(CH_3)$_2$—CH_2—O—P(O^-)(=O)—O—CH_2—Ser— ACP

β- 巯基乙胺　　泛酸

图 29-15　ACP 的结构

（四）脂肪酸合成的反应历程

现以哺乳动物为例，分步介绍脂肪酸合酶所催化的反应（图 29-16）：

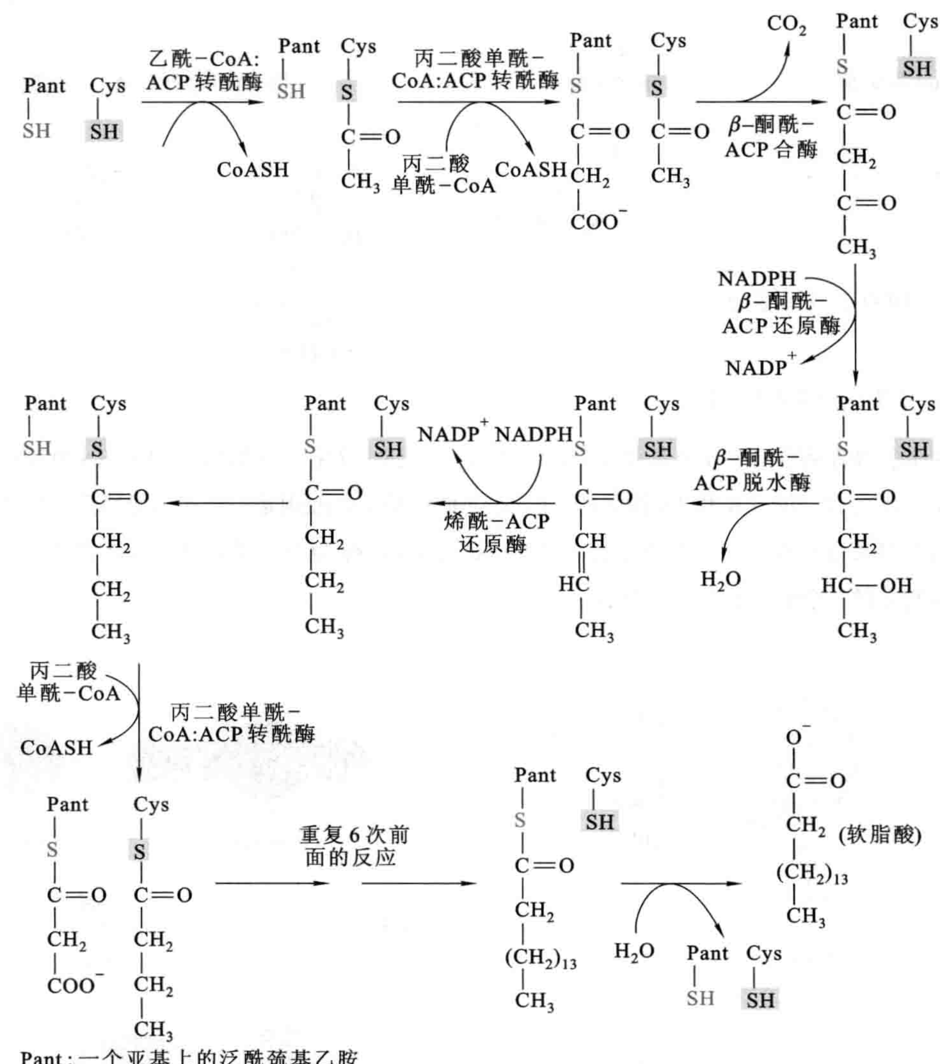


图 29-16　脂肪酸合酶催化的脂肪酸合成反应

（1）引发反应（priming reaction）　作为引物的乙酰基从 CoA 转移到脂肪酸合酶二聚体的一个亚基上，这个亚基接受乙酰基的是 KS（即 CE）上的一个 Cys 残基。该反应由 AT 催化，反应式为：

$$\text{乙酰}-S-CoA + HS-Cys-KS \longrightarrow \text{乙酰}-S-Cys-KS + CoASH$$

（2）活化的“二碳单位”的装载（loading）　反应发生在另一个亚基的 ACP 巯基上（简写为 PPant-SH），由 MT 催化，反应式为：

$$\text{丙二酸单酰}-S-CoA + HS-PPant-ACP \longrightarrow \text{丙二酸单酰}-S-PPant-ACP + CoASH$$

（3）缩合（condensation）　这是一步碳链延伸的反应，由 KS 催化，总反应式为：

$$\text{丙二酸单酰}-S-PPant-ACP + \text{乙酰}-S-Cys-KS \longrightarrow CO_2 + \text{乙酰乙酰}-S-PPant-ACP + HS-Cys-KS$$

在反应中，先是丙二酸部分发生脱羧基反应形成碳负离子中间物，随后碳负离子亲核进攻乙酰基部分的羰基碳，两个“二碳单位”因此缩合成四碳的乙酰乙酰基团。

（4）还原（reduction）　这是一步氧化还原反应，由 KR 催化，NADPH 为还原剂，产物为 D-α,β- 羟丁酰 -ACP，反应式为：

$$\text{乙酰乙酰}-S-PPant-ACP + NADPH + H^+ \longrightarrow D-\beta-\text{羟丁酰}-S-PPant-ACP + NADP^+$$

（5）脱水（dehydration）　这是一步脱水消除反应，由 DH 催化，随着位于 α- 碳原子上的氢和 β- 碳

原子上的羟基形成 H_2O，一个反式的双键在这两个碳原子之间形成，总反应式为：

$$\text{D-}\beta\text{- 羟丁酰 - S-PPant-ACP} \longrightarrow \alpha,\beta\text{- 反 - 丁烯酰 - S-PPant-ACP+H}_2\text{O}$$

(6) 再还原　这又是一步氧化还原反应，由 ER 催化，还原剂仍然是 NADPH，产物为丁酰 - S-PPant-ACP，反应式为：

$$\alpha,\beta\text{- 反 - 丁烯酰 - S-PPant-ACP+NADPH+H}^+ \longrightarrow \text{丁酰 - S-PPant-ACP}$$

到此为止，两个碳原子的乙酰基被延长为 4 个碳原子的丁酰基，下面发生的反应即是重复第 1 步—第 6 步的反应，不过第 1 步变成丁酰基从一个亚基的 ACP 移位到另一个亚基的 KS 的 Cys 残基上。显然，每重复一次，碳链即延长两个碳原子。

(7) 软脂酸的释放　当碳链延伸到 16 个碳原子并被还原成软脂酰 -S-PPant-ACP 以后，TE 将 ACP 上的软脂酰键转移给水分子，反应终产物软脂酸得以释放，反应式为：

$$\text{软脂酰 -S-PPant-ACP+H}_2\text{O} \longrightarrow \text{软脂酸 +ACP-PPant-SH}$$

综合上述所有的反应，合成 1 分子软脂酸的总反应式为：

$$\text{乙酰 -CoA + 7 丙二酸单酰 -CoA + 14NADPH + 14H}^+ \longrightarrow \text{软脂酸 + 7CO}_2\text{ + 14NADP}^+\text{ + 8HSCoA + 6H}_2\text{O}$$

$$\text{7 乙酰 -CoA + 7CO}_2\text{ + 7ATP} \longrightarrow \text{7 丙二酸单酰 -CoA + 7ADP + 7Pi + 7H}^+$$

$$\text{8 乙酰 -CoA + 14NADPH + 7H}^+\text{ + 7ATP} \longrightarrow \text{软脂酸 + 14NADP}^+\text{ + 8HSCoA + 6H}_2\text{O+ 7ADP + 7Pi}$$

(五) 脂肪酸的修饰

软脂酸是脂肪酸合酶的终产物，但它并不是机体内唯一的脂肪酸，那么机体内其他的脂肪酸是如何产生的呢？

机体内其他脂肪酸主要通过各种修饰反应产生，有延伸和去饱和两种方式，相关的酶分别称为延伸酶(elongase)和去饱和酶(desaturase)。下面就这两种修饰反应分别加以讨论。

1. 脂肪酸的延伸反应

如图 29-17 所示，真核细胞有两种脂肪酸延伸反应系统，一种存在于内质网，另一种存在于线粒体基质。前者与细胞质基质的脂肪酸合成反应相似，只是酶的组成有所改变，并且使用 CoA 代替 ACP 作为脂酰基的载体；后者可视为脂肪酸 β- 氧化的逆反应，不过延伸反应最后一步(相当于 β- 氧化的第一步氧化还原反应)的电子供体并非 $FADH_2$，而是 NADPH。

2. 脂肪酸的去饱和反应

在去饱和酶作用下，饱和脂肪酸形成不饱和脂肪酸，单不饱和脂肪酸形成多不饱和脂肪酸。如图 29-18 所示，酵母和动物细胞内的去饱和酶位于光面内质网膜上，是一种含有非血红素铁的蛋白质，它与细胞色素 b_5、细胞色素 b_5 还原酶一起完成去饱和反应。在反应中，NADH 和脂酰 -CoA 同时提供电

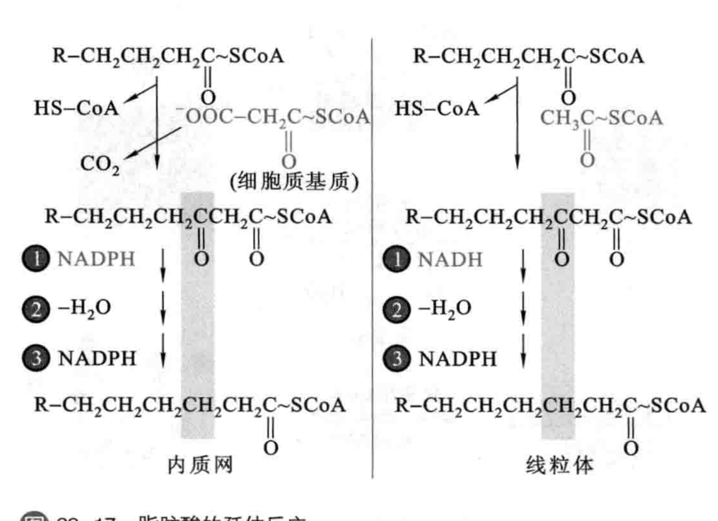


图 29-17　脂肪酸的延伸反应

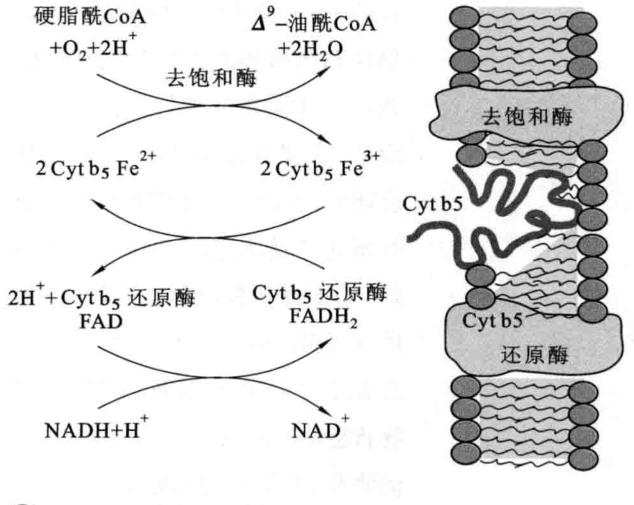


图 29-18　脂肪酸的去饱和反应

子并最终交给 O_2。植物细胞的去饱和酶是可溶性的，并不与膜结合，它作用于脂酰 -ACP，由脂酰 -ACP 和 NADPH 提供电子。

哺乳动物细胞的去饱和能力有限，不能在编号高于 9 号位 C 原子以上的位置直接引入双键，但植物细胞没有此限制（图 29-19）。在动物细胞内，像亚油酸 12 号位置的双键和 α- 亚麻酸 12 号以及 15 号位置的双键是无法引入的，只有在植物细胞内才可以引入。正因为如此，人体需要的亚油酸和 α- 亚麻酸必须从食物中获取，即这两种脂肪酸是人的必需脂肪酸。

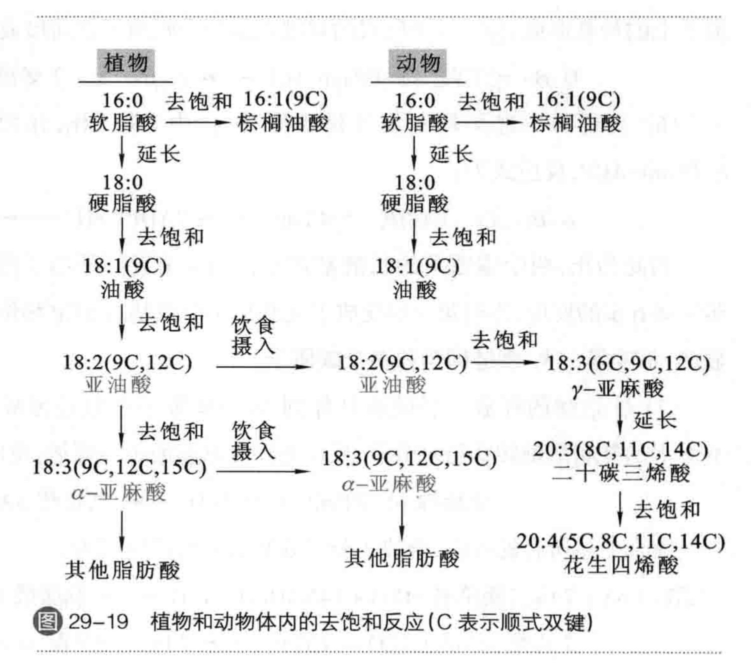


图 29-19 植物和动物体内的去饱和反应（C 表示顺式双键）

Quiz12 人体可以自己合成 18∶1(11)吗？为什么？

（六）奇数脂肪酸的合成

奇数脂肪酸除了通过偶数脂肪酸的 α- 氧化转变而成以外，还可以直接合成。合成的机制类似于偶数脂肪酸，只是使用丙酰 -CoA（propionyl-CoA）代替乙酰 -CoA 作为引物。

第三节 脂肪酸代谢的调控

一、脂肪酸分解代谢的调控

Quiz13 机体为什么没有选择脂肪酸的活化这一步作为脂肪酸分解代谢的调控位点？

脂肪酸分解代谢的主要调控位点是在脂酰 -CoA 通过线粒体内膜进入基质这一步，受到调节的限速酶是 CPT Ⅰ，丙二酸单酰 -CoA 既是该酶的抑制剂，又是脂肪酸合成的前体，其浓度是由 ACC 控制的，至于 ACC 活性的调节详见脂肪酸合成代谢的调控。

二、脂肪酸合成代谢的调控

脂肪酸合成的限速酶为 ACC。以哺乳动物为例，ACC 的调节方式有两种：一种是由别构调节引起的单体和多聚体形式的互变（图 29-20）。其中单体由四个亚基组成，无活性，多聚体是由 7~14 个单体聚合而成，有活性。在细胞能量状态较高的时候，细胞内高能荷的指示剂——柠檬酸促进单体转变为多聚体，以使细胞内多余的能量及时贮存起来。相反，脂肪酸合成的终产物软脂酰 -CoA 会以负反馈的形式，促使多聚体转变为无活性的单体。

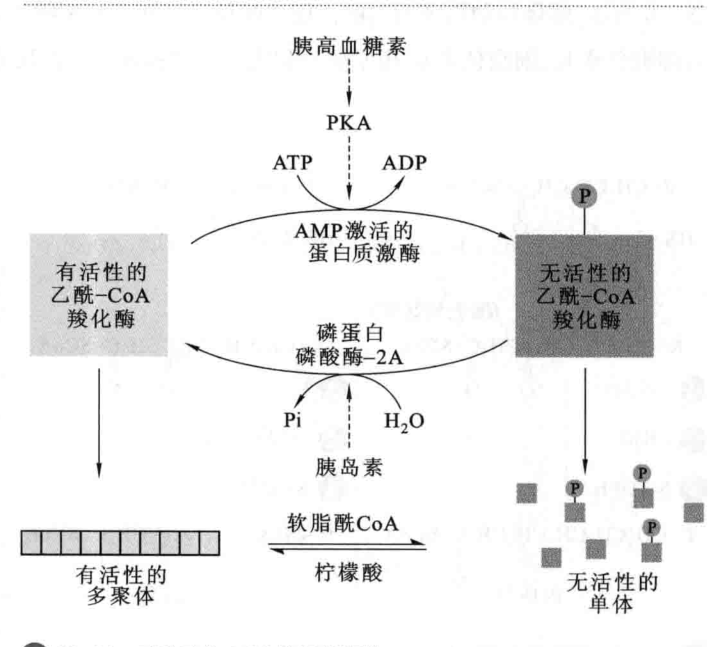


图 29-20 哺乳动物 ACC 的活性调节

调节 ACC 活性的另外一种方

式是蛋白质的磷酸化。ACC 具有磷酸化和去磷酸化两种形式,其中磷酸化为无活性形式,去磷酸化为活性形式。这种方式与第一种调节方式一致,原因在于磷酸化的乙酰 -CoA 带上了较高的负电荷,不利于有活性的多聚体形成。

促进 ACC 磷酸化的蛋白质激酶有 PKA 和一种直接受 AMP 激活的激酶(AMP-activated kinase)(参看第三十章“胆固醇代谢”),其中前者受胰高血糖素的调控(参看第十七章“激素及其受体介导的信号转导”),后者直接受 AMP 的激活。而促进 ACC 去磷酸化的酶是磷蛋白磷酸酶 -2A,该酶的活性受到胰岛素的调节。由此可见,PKA 和直接受 AMP 激活的蛋白质激酶分别在饥饿(胰高血糖素分泌加强)和细胞处于低能荷(高浓度的 AMP)状态下刺激脂肪酸的分解、抑制脂肪酸的合成以满足细胞对 ATP 的需要。

脂肪酸代谢除了通过对限速酶进行“质变”调控以外,还可以通过“量变”的方式对参与脂代谢的多个酶基因的表达进行调控。而“量变”中的一种重要方式是通过过氧化物酶体增殖激活受体作用(peroxisome proliferator-activated receptor,PPAR)进行的。

PPAR 位于细胞核内,是一类受配体结合控制的转录因子。它们最初发现可以与一类人工合成的小分子药物结合,这类小分子药物被称为过氧化物酶体增殖物,后来则发现还能与细胞内源的配体分子结合,并参与脂代谢。与 PPAR 结合的内源性细胞配体有长链多不饱和脂肪酸、氧化的脂肪酸和花生四烯酸的衍生物。在有配体的时候,PPAR 可以与视黄酸类 X 受体(retinoid X receptor,RXR)形成异源二聚体。这种异源二聚体可结合到多个参与脂代谢的基因启动子附近的过氧化物酶体增殖应答元件(peroxisome proliferator response element,PPRE)上,激活相关基因的转录。有三种 PPAR,即 PPAR-α、PPAR-β 和 PPAR-γ,但参与脂代谢的主要是 PPAR-α 和 PPAR-γ。

PPAR-α 和 PPAR-γ 在脂质代谢中分别具有不同的功能:PPAR-γ 主要存在于白色脂肪组织,与其结合的配体是 15- 脱氧 -D- 前列腺素 J_2。在生理状态下,它主要在进食后活化,可激活参与脂肪酸贮存的酶或蛋白质(如脂蛋白脂肪酶、脂肪细胞脂肪酸结合蛋白和脂酰 -CoA 合成酶)的基因表达。从而促进脂肪在白色脂肪组织中聚集;而 PPAR-α 主要分布在褐色脂肪组织和肝细胞内,与其结合的配体是长链多不饱和脂肪酸(如亚麻酸)、氧化的脂肪酸和花生四烯酸类衍生物。它在饥饿状态下活化,可激活参与脂肪酸运输以及在过氧化物酶体和线粒体内参与脂肪酸氧化的酶的基因表达,从而促进脂肪酸的动员、活化和降解,有助于降低血脂。现在许多能够结合 PPAR 的外源激动剂已成为治疗脂代谢异常的有效药物,如曲格列酮(rezulin)和文迪雅(avandia)是 PPAR-γ 的激动剂,用于治疗Ⅱ型糖尿病,而吉非罗齐(gemfibrozil)是 PPAR-α 的激动剂,用于治疗高血脂。

此外,胰岛素在肝细胞中,可以刺激脂肪酸合酶基因的表达。这表明在高血糖的条件下,机体内的脂肪酸合成加强,以便将多余的葡萄糖转化为脂肪酸贮存起来。胰岛素促进脂肪酸合酶的表达是通过上游刺激因子(upstream stimulatory factors,USFs)和胆固醇应答元件结合蛋白 -1(sterol response element binding proteins,SREBP-1)进行的;反之,多不饱和脂肪酸可通过抑制 SREBP 的产生来阻止脂肪酸合酶基因的表达。

脂瘦素是由脂肪细胞在过量的脂肪贮存的情况下产生的,其功能是调节食物摄取和脂肪代谢。它能够降低食欲,提高能量的消耗并抑制脂肪酸的合成,从而达到减肥功效。在脂肪细胞中,脂瘦素同时抑制 SREBP-1 和脂肪酸合酶基因的表达,从而抑制脂肪酸的合成。

框 29-1 生化与健康——减肥的新希望

减肥似乎是现代人道不完、谈不尽的永久话题!这次与减肥有关的话题是 ACC 这种酶。

已在哺乳动物体内发现两种形式的 ACC:一种是存在于肝细胞和脂肪细胞中的 ACC1,另一种是存在于心肌细胞和骨骼肌细胞中的 ACC2。这两种形式的 ACC 由不同的基因编码。显然,ACC1 在肝细胞和脂肪细胞中所起的作用主要是为脂肪酸合成提供活化的“二碳单位”。但是,心肌细胞和骨骼肌细胞一

般不合成脂肪酸，那为什么要表达 ACC2 呢？因为这两种细胞不能进行脂肪酸的合成，但可以进行脂肪酸的氧化，而脂肪酸氧化的限速酶是 CPT Ⅰ。调节 CPT Ⅰ活性的唯一手段是通过丙二酸单酰 -CoA 对它进行抑制。因此，ACC2 的功能最有可能是在需要抑制脂肪酸氧化的时候，为心肌细胞和骨骼肌细胞制造 CPT Ⅰ的抑制剂——丙二酸单酰 -CoA，例如，机体在处于高糖和高胰岛素的状态下。相反，当机体处于饥饿状态，或者在做剧烈运动的时候，ACC2 因磷酸化而失去活性，脂肪酸氧化变得活跃。

为了进一步确定 ACC2 的功能，Abu-Elheiga 等人在 2001 年，使用基因敲除技术得到了 *ACC2* 基因缺失的小鼠。结果发现，它们与正常小鼠的差别是：吃得多，但体内贮存的脂肪量却显著下降，只有正常小鼠的一半。在缺乏 *ACC2* 小鼠的心肌和骨骼肌中，脂肪的量分别降低了 10% 和 30%~40%，而在脂肪组织内，脂肪下降了近一半。正常小鼠多油的肝看起来有点苍白，但缺乏 *ACC2* 小鼠的肝是鲜红色的，几乎没有脂肪。正常的小鼠体内，胰岛素能刺激 ACC1 和 ACC2 的激活酶的产生，从而提高脂肪酸的合成和抑制脂肪酸的氧化。然而，缺乏 *ACC2* 的小鼠体内，脂肪酸在有胰岛素的时候照样氧化。除了脂肪含量下降以外，其余功能一切正常，如正常繁殖。

上述发现为人类寻找新的减肥药物的作用靶点提供了新的方向，若是能够得到一种 ACC2 特异性的高效抑制剂，那这种抑制剂极有可能是一种有效的减肥药物。

e29-4 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Fatty_acid_metabolism（维基百科有关脂肪酸代谢的内容）
2. http://themedicalbiochemistrypage.org/lipid-synthesis.html（完全免费的医学生物化学课程网站，有关脂肪酸代谢的内容）
3. http://web.virginia.edu/Heidi/chapter24/chp24frameset.htm（美国弗吉尼亚大学生物化学网络课程，有关糖原代谢的内容）
4. http://www.wiley.com/college/boyer/0470003790/animations/fatty_acid_metabolism/fatty_acid_metabolism.htm（内有脂肪酸代谢的动画）

参考文献：

1. Zoete V, et al. Peroxisome proliferator-activated receptor structures: ligand specificity, molecular switch and interactions with regulators. *Biochim Biophys Acta*, 1 771 (8): 915–925, 2007.
2. Berger J, Moller D E. The mechanisms of action of PPARs. *Annu Rev Med*, 53: 409–435, 2002.
3. Munday M R. Regulation of mammalian acetyl-CoA carboxylase. *Biochem Soc Trans*, 30: 1 059–1 064, 2002.
4. Pourfarzam M, et al. Mammalian mitochondrial β-oxidation. *Biochem J*, 320: 345–357, 1996.
5. Smith S. The animal fatty acid synthase: one gene, one polypeptide, seven enzymes. *FASEB J*, 8: 1 248–1 259, 1994.

第三十章　胆固醇代谢

胆固醇既是动物细胞膜的组分，又是多种生物活性物质的前体，它在生物体内的代谢转变以及转运一直是生化学家关注的焦点之一，且它与人体健康之间的关系长期以来也颇受争议。

本章将重点介绍胆固醇的生物合成、胆固醇在机体内的转运、胆固醇的代谢转变以及相关的调节机制。

第一节　胆固醇的合成

哺乳动物所有的细胞几乎都能合成胆固醇，其中合成最活跃的细胞是肝细胞（占 80%），其次是小肠上皮细胞（占 10%）和表皮细胞（占 5%）。在细胞内，合成胆固醇的场所是细胞质，其中一部分反应在细胞质基质进行，另一部分则在内质网膜上进行。

同位素示踪实验表明，胆固醇是由乙酸转变而成，但实际上合成胆固醇的直接前体是乙酰 -CoA。整个胆固醇的生物合成反应可分为 4 个阶段：

1. 3 个乙酰 -CoA →甲羟戊酸

此阶段包括三步反应（图 30-1），其中前两步与酮体合成反应的前两步相同（参看第二十九章“脂肪酸代谢”），只是反应的场所不一样。最后一步反应是不可逆的，发生在细胞质基质，由 HMG-CoA 还原酶（HMG-CoA reductase）催化，NADPH 为电子供体，主要产物为甲羟戊酸（mevalonate）。

真核细胞的 HMG-CoA 还原酶是一种膜内在蛋白，定位在内质网膜上，其跨膜结构域横跨内质网膜共达 8 次之多，活性中心位于面向细胞质基质的 C 端结构域。某些细菌也含有 HMG-CoA 还原酶，

Quiz1 从代谢的角度，肝细胞是一个“利他主义者”，就此请给出几个例子加以说明。

Quiz2 如果将乙酸分子中的羧基 C 用 ^{14}C 标记，那么最后得到的胆固醇分子中有多少个 C 被标记？

$2\,CH_3C(=O)-S-CoA$ —（乙酰乙酰-CoA硫解酶；CoASH）→ $CH_3C(=O)-CH_2-C(=O)-S-CoA$ 乙酰乙酰-CoA

乙酰乙酰-CoA + $CH_3C(=O)-S-CoA$ —（HMG-CoA 合酶；CoASH）→ $^{-}O-C(=O)-CH_2-C(OH)(CH_3)-CH_2-C(=O)-S-CoA$ β-羟-β-甲基戊二酸单酰CoA（HMG-CoA）

HMG-CoA + 2NADPH —（HMG-CoA 还原酶）→ $2NADP^+$ + CoASH + $^{-}O-C(=O)-CH_2-C(OH)(CH_3)-CH_2-CH_2-OH$ 甲羟戊酸

图 30-1　胆固醇合成第一个阶段的反应

这些细菌虽不能合成胆固醇，但可以合成甲羟戊酸来作为十一葵烯醇(undecaprenol)、维生素 K_2 和泛醌等萜类物质的前体。

Quiz3 一种柠檬酸的类似物可抑制柠檬酸裂合酶的活性，你认为这种物质在人体内能降低胆固醇吗？

HMG-CoA 还原酶是胆固醇合成的限速酶，因此它成为许多降胆固醇药物作用的理想靶标。例如，斯达汀(statin)就是一种能够有效降低血胆固醇的药物，其本质是 HMG-CoA 的类似物，在体内作为 HMG-CoA 还原酶的竞争性抑制剂起作用。

2. 甲羟戊酸→活化的异戊二烯

此阶段包括 4 步反应(图 30-2)，前两步分别由甲羟戊酸激酶(mevalonate kinase)和磷酸甲羟戊酸激酶(phosphomevalonate kinase)催化，各消耗 1 分子 ATP，甲羟戊酸经磷酸甲羟戊酸转变为 5- 焦磷酸甲羟戊酸(5-pyrophosphomevalonate)。第三步是依赖于 ATP 的脱羧反应，由焦磷酸甲羟戊酸脱羧酶(pyrophosphomevalonate decarboxylase)催化。当焦磷酸甲羟戊酸失去羧基以后，活化的异戊二烯单位——异戊二烯焦磷酸(isopentenyl pyrophosphate)随之产生。然后异戊二烯焦磷酸在一种异构酶的催化下，转变成它的异构体——二甲烯丙基焦磷酸(dimethylallyl pyrophosphate)。

活化的异戊二烯中间物的发现，为人们认识、理解一系列重要萜类物质生物合成打开了一扇窗口。许多重要的生物活性物质，如胡萝卜素、CoQ、质体醌、天然橡胶、多萜醇、两种植物激素(脱落酸和赤霉酸)、叶绿醇、四种脂溶性维生素以及古菌醚磷脂的疏水尾巴都是由它衍生而来的。

3. 6 个活化的异戊二烯单位→ 30 碳的碳氢化合物鲨烯

此阶段主要包括三步反应(图 30-3)：第一步是在二甲烯丙基转移酶(dimethylallyl transferase)的催化下，一个二甲烯丙基与一个异戊二烯焦磷酸头尾缩合成牻牛儿焦磷酸(geranyl pyrophosphate)；第二步是在牻牛儿转移酶(geranyl transferase)的催化下，牻牛儿焦磷酸与另一个异戊二烯焦磷酸头尾缩合，形成法尼焦磷酸(farnesyl pyrophosphate)；最后一步是在法尼酰转移酶(farnesyl transferase)或鲨烯合酶(squalene synthase)的催化下，两个法尼焦磷酸头头缩合，并被 NADPH 还原为鲨烯(squalene)。

法尼焦磷酸不仅是胆固醇合成的中间物，还是生物体内其他一些异戊二烯类化合物(isoprenoid)合成的前体，如多萜醇和 CoQ。此外，法尼焦磷酸还可以作为法尼基的供体，参与某些蛋白质的翻译后加工，使这些蛋白质带上高度疏水的异戊二烯单位，以便于锚定在细胞膜上，如 Ras 蛋白(参看第四十章“蛋白质的翻译后加工及其定向和分拣”)。

4. 鲨烯→胆固醇

鲨烯合成好以后，由于其不溶于水，因此需要细胞质基质中的固醇载体蛋白(sterol carrier protein，SCP)将其转运到内质网，在内质网膜上作为最后一个阶段反应的起始底物。

图 30-2 胆固醇合成的第二个阶段的反应

二甲烯丙基焦磷酸

异戊二烯焦磷酸

二甲烯丙基转移酶

PPi

牻牛儿焦磷酸

异戊二烯焦磷酸

牻牛儿转移酶

PPi

法尼基磷酸

NADPH + H^+

法尼基磷酸

法尼酰转移酶

$NADP^+$

2 PPi

鲨烯

图 30-3　胆固醇合成的第三个阶段的反应

这一个阶段的反应最为复杂，一共有 22 步，比较重要的反应有 3 步（图 30-4）：①在鲨烯单加氧酶（squalene monooxygenase）的催化下，鲨烯被氧化成 2,3- 环氧鲨烯（squalene-2,3-epoxide）；②在 2,3-环氧鲨烯羊毛固醇环化酶的催化下，环氧鲨烯被环化成羊毛固醇（lanosterol）。环化反应由一个质子启动，涉及一系列电子的转移。羊毛固醇形成以后，将连续发生 19 步反应（其中很多反应的机制还不清楚），直至形成 7- 脱氢胆固醇（7-dehydrocholestrol）；③7- 脱氢胆固醇被 NADPH 还原成胆固醇。

框 30-1　生化医药——降胆固醇药物的新靶点

2011 年 3 月 *Cell Metabolism* 杂志上一份由新南威尔士大学的 Andrew Brown 所做的研究报告表明，参与胆固醇合成的鲨烯单加氧酶（squalene mono-oxygenase，SM）是胆固醇合成的又一步重要的调控位点。

细胞内胆固醇的合成涉及 20 多种酶，人们熟悉的限速酶是 HMG-CoA 还原酶，这种酶是目前所使用的绝大多数降胆固醇药物的作用靶点，如斯达汀。HMG-CoA 还原酶催化的是胆固醇合成的第三步反应，由它合成产生的甲羟戊酸不仅参与胆固醇的合成，还参与体内其他一些重要生物分子的合成，如多萜醇和 CoQ。因此，早早的将其抑制影响的并不单单是胆固醇的合成，这就是斯达汀这一类药物产生副作用（如

鲨烯
NADPH + H^+
O_2
鲨烯单加氧酶
$NADP^+$
2,3-环氧鲨烯
羊毛固醇环化酶
H^+
羊毛固醇
19步反应
7-脱氢胆固醇
NADPH + H^+
$NADP^+$
胆固醇

图 30-4 胆固醇合成的第四个阶段的反应

肌肉痛)的原因。

与 HMG-CoA 还原酶不同的是,SM 是胆固醇合成中所特有的酶,抑制它的活性也应该能够降低胆固醇的水平,并且不会影响到体内其他物质的合成,因此不会带来不必要的副作用。

通常体内的胆固醇水平会受到严格的调控。但是,不健康的饮食习惯,以及某些遗传倾向可导致胆固醇产生的速率超过机体能够代谢它的速率,从而出现高胆固醇血脂症,并引起动脉粥样硬化和心血管疾病的发生。

现在有一种很有效的抗真菌药物叫托萘酯(tolnaftate),广泛用于浅表皮肤真菌感染和花斑癣的治疗,其作用的靶点就是真菌细胞内的 SM。这让人相信,在不久的未来,市场上一定会有以人 SM 为靶点的降胆固醇药物,并且这样的药物肯定不会存在斯达汀带来的副作用。

第二节 胆固醇的转运

已知胆固醇合成和代谢的主要场所为肝细胞。其他细胞虽然也能合成胆固醇,但合成的量通常并不能满足自身需要,这时就需要从肝细胞中获取。此外,食物中的胆固醇被小肠上皮细胞吸收以后

需要转运到肝细胞，而肝外细胞多余的胆固醇也需要转移到肝细胞进行代谢，因此，人体内存在 3 条转运胆固醇的路线（图 30–5），它们分别负责将从食物中获取的胆固醇转运到肝细胞、将肝细胞中的胆固醇转运到肝外细胞和将肝外细胞多余的胆固醇运回肝细胞。最后一条路线也被称为胆固醇的逆向转运（reverse cholesterol transport）。然而，胆固醇与其他脂质一样，其溶解性质不允许它直接在水溶性环境中进行转运，只能与脱辅基脂蛋白（apolipoprotein）组装成脂蛋白以后才能在血浆中进行转运。

Quiz4 你知道哪些食物富含胆固醇吗？

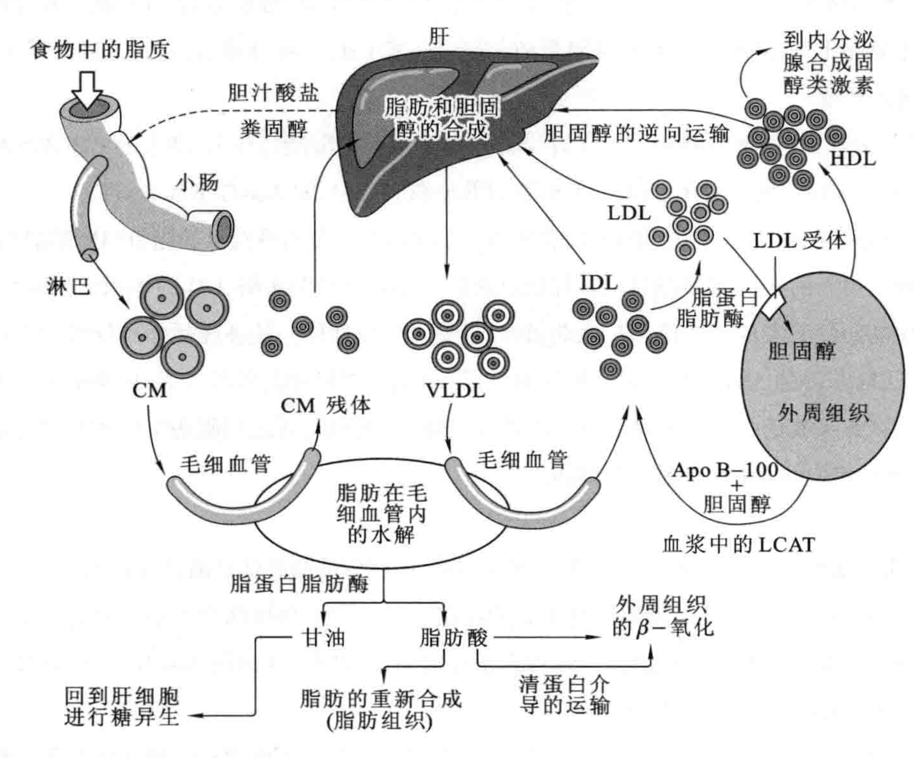


图 30–5 人体内胆固醇的转运

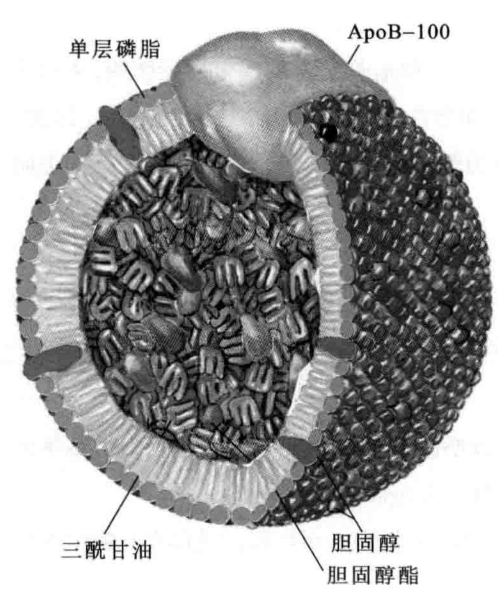


图 30–6 LDL 的结构模型（Nelson 等，2008）

按照密度从低到高的顺序，血浆中的脂蛋白可分为乳糜微粒（chylomicron，CM）、极低密度脂蛋白（very low density lipoprotein，VLDL）、中间密度脂蛋白（immediate density lipoprotein，IDL）、低密度脂蛋白（low density lipoprotein，LDL）和高密度脂蛋白（high density lipoprotein，HDL），其中 CM 来源于小肠上皮细胞，HDL 来源于肝细胞和小肠上皮细胞，VLDL 来源于肝细胞，其他两种衍生于 VLDL。这五种脂蛋白除了在密度上有差别以外，在组成成分、颗粒直径和功能上也有差别。

以 LDL 为例（图 30–6），一个典型的脂蛋白的表面有一单层的磷脂，脱辅基脂蛋白和少量的游离胆固醇分子镶嵌在其中，其内部是由脂肪、胆固醇酯通过疏水键和范德华力组成的疏水核心。

e30–1 5 种脂蛋白的主要性质

脱辅基脂蛋白就是载脂蛋白，其生理功能包括四个方面：①与脂质组装为脂蛋白，帮助脂质的转运；②作为配体，识别和结合特定细胞膜上的受体，参与受体介导的脂蛋白的内吞作用（receptor–mediated endocytosis）。例如，Apo B–100 为 LDL 受体的配体；③激活或抑制参与脂代谢的某种酶的活性。例如，CM 表面的 Apo C–Ⅱ可激活血管壁上的脂蛋白脂肪酶的活性（lipoprotein lipase，LPL），而 HDL 表面的 Apo A–Ⅰ是血清内卵磷脂，它是胆固醇脂酰基转移酶（lecithin-cholesterol acyl transferase，LCAT）的激活剂；④诱导脂蛋白残体的清除。例如，Apo E 可诱导血

液中 VLDL 和 CM 残体的清除。

已在多数载脂蛋白上发现两亲螺旋(参看第二章“蛋白质的结构”),载脂蛋白借助这种螺旋的疏水一面与脂质疏水的尾巴结合,而亲水的一面则暴露在水相之中。现就 5 种脂蛋白的结构与功能作进一步的说明。

1. CM

CM 在 5 种脂蛋白中体积最大,密度最轻。食物中的脂被消化、吸收以后,在小肠上皮细胞内的内质网和高尔基体上又重新合成,并与载脂蛋白一起组装成 CM。CM 分泌后,先进入淋巴系统,然后通过胸导管进入血液。

CM 的主要成分是脂肪(95%),还含有少量胆固醇和胆固醇酯(5%)。新生的 CM 含有载脂蛋白 Apo B-48,进入循环系统以后还可以从 HDL 获得其他载脂蛋白,如 Apo C-Ⅱ和 Apo E。

CM 一旦获得 Apo C-Ⅱ,其上的脂肪就会因 Apo C-Ⅱ激活毛细血管壁上的脂蛋白脂肪酶而被水解。脂肪水解释放出来的 FFA 和单酰甘油被体细胞吸收。随着脂肪的水解,CM 的体积不断减小,而内部脂质核心中的胆固醇酯的相对比例却逐渐增大,最终蜕变成残体。残体经受体介导的内吞作用被肝细胞吸收。CM 获得的 Apo E 作为配体起作用。当 CM 残体被肝细胞吸收以后,原来的胆固醇便进入了肝细胞。因此,CM 的主要功能是将食物中被小肠吸收的脂肪,转运到脂肪组织和其他体细胞作为燃料,同时将食物中的胆固醇转移到肝细胞。

2. VLDL

VLDL 由肝细胞装配和分泌,在其装配过程中,内质网腔中的微粒体三酰甘油转移蛋白(microsomal triglyceride transfer protein,MTP)起着十分重要的作用。MTP 是一种由两个大小不同的亚基组成的异源二聚体蛋白。其大亚基能够促进脂质转移到正在糙面内质网上合成的 Apo B-100 的附近,而小亚基具有蛋白质二硫化物异构酶的活性。

VLDL 分泌后直接进入血液,其核心含有高比例的脂肪(55%),也含有一定比例的胆固醇和胆固醇酯(25%)。新生的 VLDL 含有载脂蛋白 Apo B-100,进入循环后也可以从 HDL 得到其他载脂蛋白,如 Apo C-Ⅱ和 Apo E。

和 CM 一样,在循环中,位于 VLDL 表面的 Apo C-Ⅱ激活毛细血管壁上的脂蛋白脂肪酶,从而导致其脂肪水解。随着脂肪的丢失,VLDL 的体积逐渐减小,而密度增大,它经 IDL 转变为 LDL。因此,VLDL 的功能是将内源的脂肪从肝细胞转运到肝外组织作为燃料,而其中的胆固醇和胆固醇酯的去向则由 LDL 决定。

3. IDL

IDL 由 VLDL 转变而来,其脂肪的比例降低到 20%,而胆固醇的比例则上升到 40%。血液中的 IDL 约有一半被肝细胞吸收,另一半则在失去更多的脂肪后转变成 LDL。

4. LDL

LDL 由 IDL 转变而来,其中的脂肪比例已降到 5%,胆固醇的比例则升至 50%。此外,LDL 将绝大多数 Apo C-Ⅱ和 Apo E 返还给了 HDL,因此它主要的载脂蛋白为 Apo B-100。

由于绝大多数细胞的质膜(包括肝细胞)含有 LDL 的受体,因此 LDL 可以通过受体介导的内吞作用被肝外细胞吸收,还有少数 LDL 可以被肝细胞吸收。

LDL 受体的突变会导致 LDL 无法被细胞吸收利用,导致血浆中的 LDL 水平严重超标,从而加速或诱发动脉粥样硬化。由此可见,LDL 的主要功能是将肝细胞中的胆固醇转运给肝外细胞,以满足肝外细胞对胆固醇的需求。有一种家族性高胆固醇血脂症(familial hypercholesterolemia)就是因为 LDL 受体发生了突变,使其不能固定在膜上,而是直接释放到血液之中,由于患者体内的 LDL 无法被细胞吸收,血浆胆固醇水平会越来越高,出生后不久就会很高。随年龄的增长,患者身体许多部位都会出现皮肤黄色瘤和肌位黄色瘤,大多数患者在 40 岁以前就有严重而广泛的动脉粥样硬化,严重者 3 岁时就死于心肌梗死。

LDL中的胆固醇通常称为“坏胆固醇”(bad cholesterol),就是因为它与动脉粥样硬化(atherosclerosis)的发生有关。LDL与动脉粥样硬化的关系如下:当血管内皮因高血压或吸烟等因素受损以后,LDL透过血管壁而沉积在其中,随着时间的推移,LDL被氧化。氧化的LDL引来巨噬细胞,并被“吞并”。巨噬细胞因吞噬LDL而充满脂类,形如泡沫,死亡后在血管壁释放出大量的脂。随后在细胞因子和生长因子的作用下,平滑肌细胞覆盖在脂质上形成胶原状的帽子。帽状结构会变得越来越大,而变大的帽状结构又吸引新的巨噬细胞前来对其降解。帽状结构因此破裂并暴露出其内部的胶原和脂质,最终导致血小板在此处的聚集和凝血,血管有随时堵塞破裂的危险。如果这种情况发生在冠状动脉,就极可能导致心肌梗死;如果发生在脑部,就可能导致中风。抗氧化剂,如维生素E和C能够抑制LDL的氧化,因此具有预防心肌梗死和中风的功效。

5. HDL

HDL由肝细胞和小肠上皮细胞装配和分泌,其脂肪和胆固醇的含量分别为5%和20%。HDL富含各种载脂蛋白,它既可以为其他脂蛋白提供载脂蛋白,也可以从其他脂蛋白那里接受载脂蛋白。

HDL的重要功能是参与胆固醇的逆向转运。这种转运首先需要将胆固醇从外周细胞定向转移到HDL颗粒上,此过程与ABC1转运体和LACT有关。ABC1转运体位于细胞膜上,它通过水解ATP驱动胆固醇从膜的内层转移到外层。一旦胆固醇到达膜的外层,就可以通过自由扩散进入HDL。显然,如果进入HDL的胆固醇不作任何改变,HDL上的胆固醇也就可能反过来进入外周细胞膜,不过这种情形很难发生,原因是HDL中的Apo A-Ⅰ能够激活血液中LCAT的活性。LCAT被激活后,催化颗粒表面的卵磷脂2号位的脂酰基使胆固醇酯化。而HDL上游离的胆固醇一旦被酯化,就会因失去两性性质进入HDL的疏水核心,这就在外周细胞膜和HDL表面之间,形成一种有利于胆固醇向HDL转移的浓度梯度,促使膜上多余的胆固醇转移到HDL。此外,HDL还可以从死亡的细胞或其他脂蛋白得到胆固醇。

HDL在得到胆固醇以后,既可以通过脂交换将一部分胆固醇转移给正在循环的VLDL,又可以通过受体介导的内吞作用被肝细胞吸收。HDL与VLDL之间发生的脂交换由胆固醇酯转移蛋白(cholesterol ester transfer protein,CETP)催化。CETP催化HDL内的胆固醇酯(CE)与VLDL内的脂肪进行交换。进入肝细胞的胆固醇可被降解成胆汁酸后排出体外,也可以重新装配到VLDL之中,进入循环,在肝外组织重新分配。因此,HDL被视为血液中的清道夫,它上面的胆固醇被称为“好胆固醇”(good cholesterol)。临床实验证明,一个人血液内HDL的量越高,其动脉粥样硬化的发生率就越低。

HDL上胆固醇的清除还可以由清道夫受体-B1(scavenger receptor B1,SR-B1)介导。SR-B1存在于许多类型的细胞的质膜上,但它并不能诱发受体介导的内吞作用。一旦与SR-B1结合,HDL内的胆固醇和胆固醇酯就被转移到细胞。当丧失了胆固醇和胆固醇酯以后,HDL与SR-B1解离,重新进入循环。在某些以胆固醇作为生物合成前体的细胞上,SR-B1数目可以发生上调,以满足这些细胞对胆固醇的大量需要。

现在有人发现了一种治疗高胆固醇血脂症的新方法,就是合成一种人工的HDL,并将其导入到体内,从而为机体提供更多的“好胆固醇”。人工合成的HDL在大小上与天然HDL差不多,表面模拟了天然HDL表面的组成和性质,其核心是金纳米颗粒,包被两层脂,且其表面插有天然HDL特有的一种主要载脂蛋白——ApoA1。

第三节 胆固醇的代谢转变与调节

一、胆固醇的代谢转变

机体内的胆固醇代谢是非常活跃的,它既可以转变成胆固醇酯以便于转运和贮存,又可以作为多

种物质的前体，参与这些物质的生物合成。

1. 胆固醇的酯化

催化胆固醇酯化反应的酶有胆固醇脂酰基转移酶（acyl-CoA：cholesterol acyl transferase，ACAT）和LCAT 两种，前者位于胞内，后者在胞外，存在于血液之中，需要 Apo A-I 的激活。这两种酶催化的反应式如下：

$$\text{脂酰 -CoA+ 胆固醇} \xrightarrow{\text{ACAT}} \text{胆固醇酯 + 脂肪酸}$$

$$\text{卵磷脂 + 胆固醇} \xrightarrow{\text{LCAT}} \text{胆固醇酯 + 溶血卵磷脂}$$

2. 胆固醇→胆汁酸

胆汁酸作为天然的去污剂，在脂肪和其他脂类的消化和吸收中发挥着重要作用（参看第二十一章生物大分子的消化和吸收），其合成的前体是胆固醇，合成和贮存的场所分别是肝细胞和胆囊。胆固醇在从体内排出之前必须进入肝组织，并转化为胆汁酸或以中性固醇的形式随胆汁分泌到消化道，在消化道约有 5% 排出体外，95% 被重吸收经门脉回到肝脏，因此胆汁酸可视为体内清除胆固醇的唯一通道。

胆汁酸有初级胆汁酸和次级胆汁酸之分。前者在肝细胞内形成，包括胆酸（cholic acid）、鹅脱氧胆酸以及它们与甘氨酸或牛磺酸（taurine）形成的结合物；后者是在初级胆汁酸被分泌到肠道以后，经肠道细菌进一步作用转化而来，包括 7- 脱氧胆酸和石胆酸（lithocholic acid）。

e30-2 胆酸和甘氨胆酸的形成

3. 7- 脱氢胆固醇→维生素 D

皮下组织中的 7- 脱氢胆固醇在中波紫外光（UVB）作用下，发生光解，形成维生素 D_3。但若不在皮肤细胞，7- 脱氢胆固醇就可能转变为光固醇（lumisterol）和速固醇（tachysterol）。

e30-3 固醇类激素的形成

胆固醇在体内除了可以发生酯化、转变成胆汁酸和维生素 D 以外，还可以作为合成固醇类激素的前体。

二、胆固醇代谢的调节

胆固醇合成的限速酶是 HMG-CoA 还原酶，该酶为胆固醇代谢调控的主要位点。调节的方式包括“蛋白质的可逆磷酸化”、酶的降解和酶基因的表达调控这三种方式，现分别加以讨论。

1. HMG-CoA 还原酶的可逆磷酸化

如图 30-7 所示，HMG-CoA 还原酶具有磷酸化和去磷酸化两种形式，其中去磷酸化形式才具有活性。调节磷酸化和去磷酸化反应的酶分别是 AMP- 激活的蛋白质激酶（AMP-activated protein kinase，AMPK）和HMG-CoA还原酶磷酸酶。

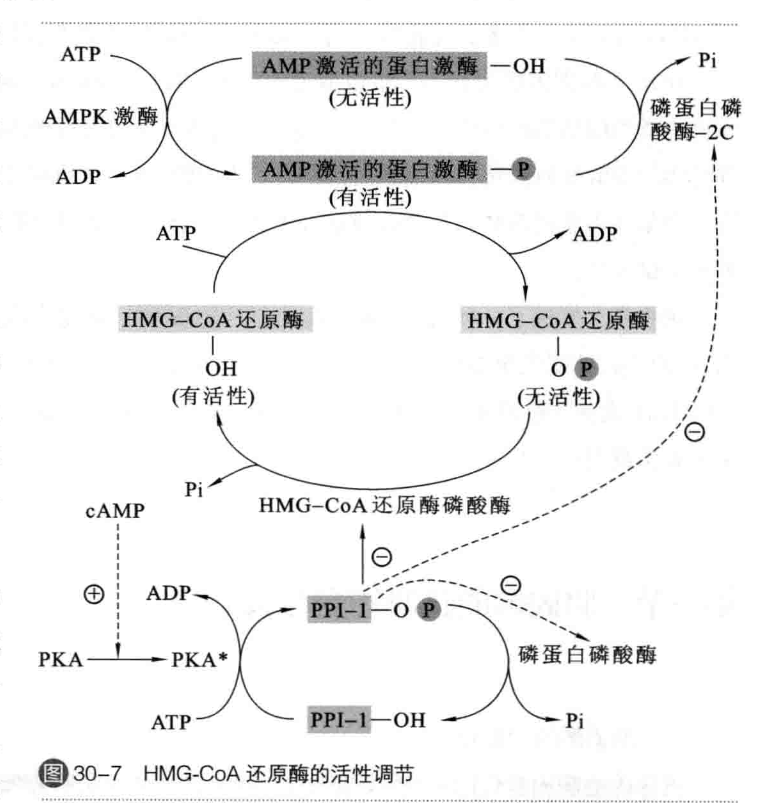


图 30-7 HMG-CoA 还原酶的活性调节

AMPK 作为机体代谢的一种总开关，广泛存在于各种真核生物中，在细胞内能实时探测和“解读”细胞的能量状态，其激活后产生的净效应是抑制合成代谢，同时刺激分解代谢。哺乳动物的 AMPK 由 α、β 和 γ 亚基组成，其中 α 亚基为催化亚基。γ 亚基有两个 AMP 结合位点，其功能是灵敏探测细胞内 AMP和ATP 的比

率。AMP 的结合具有正协同效应，即第一个 AMP 的结合使得第二个 AMP 的结合变得容易。如果细胞的能量状态低，即 AMP 水平高，那 AMP 就与 γ 亚基上的结合位点结合。一旦结合了两个 AMP，γ 亚基的构象就会发生变化，从而使 α 亚基上的催化结构域得以暴露。暴露出来的催化结构域受上游 AMPK 激酶的催化，其 Thr172 发生磷酸化修饰而使 AMPK 激活；如果细胞的能量状态高，即 AMP 水平低，这时 γ 亚基就无 AMP 结合，这使得磷酸化的 Thr172 成为磷蛋白磷酸酶 -2C 底物。Thr172 处于去磷酸化状态，AMPK 的活性因此被抑制。

HMG-CoA 还原酶磷酸酶的活性则由磷蛋白磷酸酶抑制剂 -1（PPI-1）（参看第二十七章“糖原代谢”）调节。PPI-1 在 PKA 的催化下被磷酸化以后，不仅能抑制 HMG-CoA 还原酶磷酸酶的活性，还可以抑制磷蛋白磷酸酶 -2C 的活性，这两种磷酸酶的同时抑制可以使 HMG-CoA 还原酶持续处于无活性的磷酸化状态。

Quiz5 人体内有哪些限速酶的磷酸化形式没有活性？它们各自参与何种代谢途径？

2. HMG-CoA 还原酶的降解

当细胞内的胆固醇、被氧化的胆固醇衍生物、甲羟戊酸、法尼醇（farnesol）或法尼焦磷酸过多时，HMG-CoA 还原酶的降解被激活，其机制与 HMG-CoA 还原酶本身的结构有一定的关系。有证据表明，HMG-CoA 还原酶的跨膜区含有一种固醇感应结构域，这种特殊的结构域能直接检测到细胞内胆固醇或其他一些参与胆固醇合成的中间代谢物浓度的变化。当细胞内的胆固醇或甲羟戊酸等物质浓度较高时，它们能与上述结构域结合而诱导酶的构象发生变化，使 HMG-CoA 还原酶更容易被细胞内依赖于泛素的蛋白酶体降解系统识别并降解（参看第三十九章“蛋白质的生物合成及其在细胞内的降解”）。

3. HMG-CoA 还原酶基因表达的调控

HMG-CoA 还原酶的基因表达需要一种叫固醇调节元件结合蛋白 -2（sterol regulatory element binding protein-2，SREBP-2）的转录因子，此转录因子能对细胞内固醇的浓度变化做出反应。当细胞内固醇浓度低的时候，SREBP-2 从内质网膜上的前体蛋白中被切割释放出来，并移位到细胞核作为转录因子刺激 HMG-CoA 还原酶和其他一些参与胆固醇合成的酶基因的表达。

Quiz6 利用你已学过的知识，试回答如何能够降低体内的胆固醇水平？

有三种蛋白质参与 SREBP-2 的激活释放（图 30-8）：一种是 SREBP 切割激活蛋白（SREBP cleavage-activating protein，SCAP）。SCAP 是一种内质网膜整合蛋白，其跨膜部分含有与 HMG-CoA 还原酶同源的固醇感应结构域。固醇与此结构域的结合可抑制 SCAP 的活性；另外两种是整合在高尔基体膜上的位点 1 蛋白酶（site one protease，S1P）和位点 2 蛋白酶（S2P）。具体过程是：当细胞内固醇浓度低的时候，SCAP 与 SREBP-2 的前体蛋白结合并促进其转运到高尔基体。一旦 SREBP-2 的前体到达高尔基体，早已存在于此的 S1P 和 S2P 便对其进行依次切割。其中 S1P 的切点在前体位于腔内的

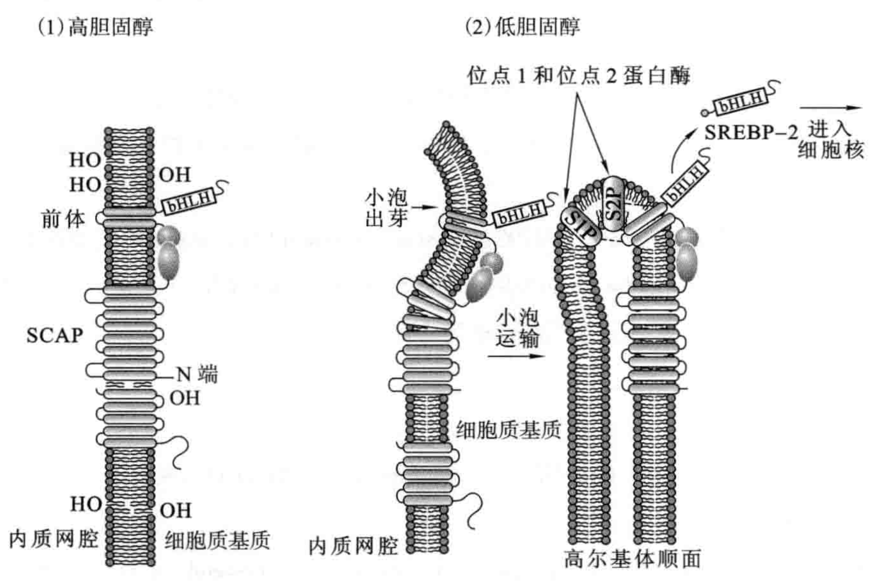


图 30-8 HMG-CoA 还原酶基因表达的调节机制（OH 为胆固醇分子上的羟基）（Lodish 等，2007）

肽段，S2P 的切点位于前体的跨膜区。在两种蛋白酶的作用下，SREBP-2 从前体的 N 端结构域释放到细胞质基质，并进入细胞核，刺激 HMG-CoA 还原酶基因的表达。

科学故事——LDL 受体的发现

家族性高胆固醇血症（familial hypercholesterolemia，FH）是一种遗传性疾病。此病患者的血清胆固醇浓度异常高，患者往往还没有成年，就会得心脏病。

1972 年，Michael Brown 和 Joseph Goldstein 在对 FH 进行系统研究的基础上指出，FH 患者胆固醇浓度过高的原因是体内调节胆固醇合成的机制有缺陷。他们的证据是：如果将 LDL 加到正常人成纤维细胞的培养基中，细胞内胆固醇合成的限速酶——HMG-CoA 还原酶的活性就会受到抑制；如果将 LDL 加到 FH 患者的成纤维细胞的培养基中，细胞内 HMG-CoA 还原酶活性就不受影响。下面的实验表明：HMG-CoA 还原酶调节的异常并不是其基因突变造成的，而可能是 FH 细胞不能从 LDL 中吸收胆固醇引起的。

1974 年，Brown 和 Goldstein 将 ^{125}I 标记的 LDL 加到正常人的成纤维细胞培养基中，然后测定在不同的保温时间内结合到细胞上的放射性强度。结果发现，和正常人成纤维细胞结合的放射性标记的 LDL 量与保温时间正相关，而且，如果同时加入过量的非放射性同位素标记的 LDL，就会减少与细胞结合的同位素标记的 LDL 的量。此外，他们还发现，在培养基中加入其他脂蛋白，却不会干扰 LDL 的结合。上述结果表明：LDL 是与细胞表面数目有限的特异性结合位点结合的。与正常人成纤维细胞和 LDL 结合实验不同的是，FH 患者的成纤维细胞不能结合同位素标记的 LDL。这样的实验结果似乎说明，正常人成纤维细胞具有一种特异性的 LDL 受体，而 FH 患者缺乏这种受体或者这种受体有缺陷。

进一步的实验证明，LDL 与正常人成纤维细胞的细胞膜结合，这意味着 LDL 受体属于细胞表面受体。他们还观察到，与细胞表面结合的 LDL 能快速地进入细胞内部，在溶酶体被降解，并释放出游离的胆固醇。后来，他们与 Richard Anderson 合作，发现 LDL 的内在化是在有被小窝区域内通过内吞作用进行的。

Brown 和 Goldstein 的工作不仅揭示了胆固醇代谢调节的主要机制，还为胆固醇代谢异常引起的疾病的治疗指明了方向。Brown 和 Goldstein 曾经帮助治疗一个叫 Stormie Jones 的 FH 病人，她从父母那里继承了有缺陷的 LDL 受体基因，其血清胆固醇含量约为 1 200 mg，几乎是正常人的 5 倍。Stormie 6 岁就得了心脏病，不得不同时接受心脏和肝脏的双器官移植手术，成为世界上第一例在一次手术中移植两种器官的人。移植的心脏取代她受损的心脏，移植的肝脏则具有正常数目的 LDL 受体。她后来因丙肝进行过第二次肝脏移植，于 13 岁去世，可能死于心脏感染。

e30-4 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Cholesterol（维基百科有关胆固醇代谢的内容）
2. http://themedicalbiochemistrypage.org/cholesterol.html（完全免费的医学生物化学课程网站，有关胆固醇代谢的内容）
3. http://www.wiley.com/college/boyer/0470003790/animations/cholesterol/cholesterol.htm（胆固醇代谢的动画）
4. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb2/part1/cholesterol.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关胆固醇代谢的内容）

参考文献：

1. Mirkin C A，et al. Templated Spherical High Density Lipoprotein Nanoparticles. *J Am Chem Soc*，131(4)：1 384-1 385，2009.
2. Brown M S，Goldstein J L. A proteolytic pathway that controls the cholesterol content of membranes，cells，and blood. *Proc Natl Acad Sci USA*，96：11 041-11 048，1999.

第三十一章 氨基酸代谢

氨基酸的主要功能是作为多肽和蛋白质的组成单位,除此以外,还可以作为多种生物活性物质合成的前体。其碳骨架则可以充当细胞的能源,在人体内可作为糖异生或酮体合成的原料,但其释放出来的氨基一般对生物体来说是有毒的,因此需要专门的解毒机制对其进行解毒,例如在人体内将其转变成尿素。

本章将重点介绍氨基酸的分解代谢以及解除氨毒的几种机制,而氨基酸及其衍生物的合成代谢仅做简单介绍。

第一节 氨基酸的分解

在结构上,氨基酸可分为氨基和碳骨架两个部分,因此,其分解代谢也可简单地分为氨基代谢和碳骨架的代谢,现就这两个部分分别加以讨论。

一、氨基的代谢

氨基的代谢有脱氨基、转氨基和联合脱氨基三种方式,其中第三种方式是前两种方式的组合。

(一) 脱氨基反应

$$\text{氨基酸} + FMN + H_2O \xrightarrow{\text{L-氨基酸氧化酶}} \alpha\text{-酮酸} + FMNH_2 + NH_4^+$$

$$FMNH_2 + O_2 \longrightarrow FMN + H_2O_2$$

$$2H_2O_2 \xrightarrow{\text{过氧化氢酶}} 2H_2O + O_2$$

图 31-1 氧化脱氨基反应

参与脱氨基反应的主要有三种酶:L-氨基酸氧化酶、D-氨基酸氧化酶和谷氨酸脱氢酶(glutamate dehydrogenase)。前两种属于黄素蛋白,分别作用于L-氨基酸和D-氨基酸。第一种酶以FMN或FAD作为辅基,第二种酶只能以FAD作为辅基(图 31-1)。在反应中,氨基酸转变为相应的α-酮酸,氨基转变为NH_4^+,脱下的氢经过FMN或FAD直接与O_2结合形成H_2O_2。有毒的H_2O_2可被细胞内的过氧化氢酶迅速分解为H_2O和O_2。

然而,L-氨基酸氧化酶在生物体内的分布并不普遍,其最适pH也远离生理pH,D-氨基酸氧化酶只作用于并不常见的D-氨基酸,因此,这两种酶并不是参与氨基酸氧化脱氨基作用的主要酶。参与氨基酸氧化脱氨基的主要酶是谷氨酸脱氢酶。在真核细胞中,它位于线粒体基质,其催化的反应属于可逆反应,反应式为:

$$\text{谷氨酸} + NAD(P)^+ + H_2O \underset{\text{ATP/GTP 促进逆反应}}{\overset{\text{ADP/GDP 激活正反应}}{\rightleftharpoons}} \alpha\text{-酮戊二酸} + NH_4^+ + NAD(P)H + H^+$$

反应的辅酶既可以是辅酶Ⅰ,又可以是辅酶Ⅱ。该酶的活性是受到严格调控的:当细胞的能量状态低的时候,细胞贫能的指示剂(ADP或GDP)将刺激正反应,即氨基酸的氧化脱氨基,此时若反应以NAD^+为电子受体,生成的NADH就进入呼吸链,若以$NADP^+$为电子受体,生成的NADPH就可作为生物合成的还原剂;当细胞的能量状态高的时候,细胞富能的指示剂(ATP或GTP)则促进逆反应,有利于谷氨酸的合成。

哺乳动物体内的谷氨酸脱氢酶还受到一种特殊的共价修饰的调节。这种共价修饰是ADP-核糖基化,催化此反应的是一种叫SIRT4的蛋白质,ADP-核糖基的供体是辅酶Ⅰ。ADP-核糖基化可抑制

谷氨酸脱氢酶的活性。这种调节机制在胰岛 β 细胞中可用来调节胰岛素的分泌，因为谷氨酸脱氢酶的 ADP- 核糖基化可下调受氨基酸刺激的胰岛素分泌。

Quiz1 试预测 SIRT4 缺失的小鼠的胰岛素的分泌有何变化?

Ser 和 Thr 的侧链上都含有容易脱去的羟基，因此，这两种氨基酸可以在相应的脱水酶(dehydratase)的催化下，以磷酸吡哆醛为辅基，直接发生脱氨基反应：

$$\text{Ser} \longrightarrow \text{丙酮酸} + NH_4^+ \qquad \text{Thr} \longrightarrow \alpha\text{- 氨基 -}\beta\text{- 羰基丁酸} \longrightarrow \text{丙酮酸} + NH_4^+$$

(二) 转氨基反应

转氨基反应是指一种氨基酸的氨基被转移到一种 α- 酮酸的羰基上，形成一种新的氨基酸和一种新的 α- 酮酸的过程。催化此反应的酶为转氨酶(aminotransferase 或 transaminase)。在反应中，磷酸吡哆醛为酶提供亲电基团参与催化(参看第十章“酶的催化机理”有关共价催化的内容。)转氨反应是完全可逆的，因此它既参与氨基酸的降解，又参与氨基酸的合成。

转氨酶有多种，但是每一种转氨酶都需要磷酸吡哆醛作为辅酶，而且绝大多数转氨酶以 Glu 作为氨基的供体，或者以 α- 酮戊二酸作为氨基的受体(图 31-2)。最常见的转氨酶有两种，即谷丙转氨酶(glutamate:pyruvate transaminase，GPT) 和谷草转氨酶(glutamate:oxaloacetate transaminase，GOT)，其中 GPT 是肝细胞内最活跃的酶之一。受损的、快要死的或已经死亡的肝组织会大量释放 GPT 进入血液，导致血清中该酶活性迅速升高，因此，临床上测定血清中 GPT 活性已成为诊断肝功能是否异常的一种常规方法。

图 31-2 转氨基反应的机理

在饥饿的时候，Ala 从肌肉组织释放到血液中，并通过丙氨酸循环(参见第二十五章“糖异生”)在肝细胞内进行糖异生，以维持血糖浓度，GPT 在此过程中也起着十分重要的作用。

GOT 也是肝细胞内一种极为活跃的酶，它存在于细胞质基质和线粒体，其重要性在于维持细胞内两大氨基酸库(amino acid pool)，即谷氨酸库和天冬氨酸库之间的平衡(图 31-3)。Glu 和 Asp 均参与尿素循环，为氨的解毒和氮的排泄所必需，因此氮原子在这两种氨基酸库之间的自由流动是细胞内氮正常代谢的关键。

由 GOT 所催化的转氨反应，不论是在细胞质基质，还是在线粒体都接近于平衡，同时 GOT 还是苹果酸 - 天冬氨酸穿梭系统中必不可少的组分(参看第二十二章“糖酵解”)，参与将细胞质基质的 NADH 转换成线粒体基质中的 NADH，以使细胞质基质的 NAD^+ 得以再生。

基于与 GPT 同样的原因，GOT 活性的变化也可以作为诊断肝功能是否异常的一种指标。

然而，并非所有的氨基酸都可以参与转氨反应，Thr、Pro 和 Lys 就是例外。

(三) 联合脱氨基反应

这是由转氨酶和谷氨酸脱氢酶组合在一起的脱氨基反应。两种酶作用的次序是：先是在转氨酶

图 31–3　谷草转氨酶催化的转氨基反应

催化下，一种氨基酸的氨基被转移到 α– 酮戊二酸的羰基上形成 Glu，然后 Glu 在谷氨酸脱氢酶催化下，发生氧化脱氨基反应，产生 α– 酮戊二酸、NH_4^+ 和 NAD(P)H。

联合脱氨基作用在氨基的最终代谢中起着重要的作用。据估计，被摄入体内的蛋白质约有 75% 是通过这种方式进行氨基代谢的。

当氨基通过氧化脱氨基或联合脱氨基反应以后，形成的氨在绝大多数生物体内还要进一步进行代谢。代谢的主要目的是解除氨的毒害。

二、氨的进一步代谢转变

氨对生物体是有毒的，氨在遇到水时很容易转变成毒性的铵盐。例如，当血液中铵盐的浓度超过 0.25 mmol/L 的时候，人体将会出现呕吐、抽搐和昏迷等症状甚至死亡。不同种类的生物处理氨的手段或者说在解除氨毒的方式上是不一样的，一般说来与一种生物对水的可得性有关，概括起来有以下几种方式：

（一）直接排出体外

水生动物（如鱼类）始终生活在水中，它们不需要对氨做任何处理，只需要将氨通过鳃直接排出体外。对于两栖动物来说，它们在蝌蚪期也是直接排氨的，但到了成体以后，生活的环境发生了根本性的改变，这时就需要将氨转变为尿素再排出体外。

（二）转变为酰胺

将氨转变为两种酰胺氨基酸中的酰胺基团是生物解除氨毒的另外一种手段，它也是细胞合成酰胺氨基酸的途径。

Asn 和 Gln 的合成反应都消耗 ATP，但消耗的方式有别（图 31–4 和图 31–5）。前者是 ATP 变成了 AMP 和 PPi，这与脂肪酸的活化反应相似（参看第二十九章“脂肪酸代谢”）。

植物通常利用 Asn 来解除氨毒，而高等动物的大部分细胞则利用 Gln 来暂时解除氨毒。在很多情况下，Gln 作为氨的载体，动物可以通过它将一种组织中产生的氨转移到另外一种组织，但最后一般都汇总到肝细胞。此外，很多生物正是利用 Gln 中的酰胺 N 作为多种含 N 物质合成的氮源（如核苷酸、某些氨基酸和氨基糖的合成）。

Quiz2　从 Asp 合成 Asn 与从 Glu 合成 Gln 消耗的 ATP 数目是一样吗？

(1)

$$NH_4^+ + \text{Asp} + ATP \xrightarrow[\text{天冬酰胺合成酶}]{Mg^{2+}} AMP + PP_i + \text{Asn}$$

(2)

$$\text{Asp} + ATP \xrightarrow[-PP_i]{(a)} \beta\text{-天冬氨酰AMP} \xrightarrow[-AMP]{(b)\ +NH_4^+} \text{Asn}$$

图 31-4　Asn 生物合成及其在植物氨解毒中的作用

(1)

$$NH_4^+ + \text{Glu} + ATP \xrightarrow[\text{谷氨酰胺合成酶}]{Mg^{2+}} ADP + P_i + \text{Gln}$$

(2)

$$\text{Glu} + ATP \xrightarrow[-ADP]{(a)} \gamma\text{-谷氨酰磷酸} \xrightarrow[-HPO_4^{2-}]{(b)\ +NH_4^+} \text{Gln}$$

图 31-5　Gln 的生物合成及其在动物氨解毒中的作用

正因为 Gln 在生物合成反应中起着枢纽作用，所以催化 Glu 转变为 Gln 的酶即谷氨酰胺合成酶（glutamine synthetase，GS）的活性受到严格的调控。以大肠杆菌的 GS 为例，该酶是一种十二聚体（dodecamer）蛋白，共受到 9 种物质的反馈抑制（图 31-6）。

这 9 种反馈抑制剂要么是 Gln 作为前体参与的合成代谢的终产物（如 Trp、His、6- 磷酸葡糖胺、氨

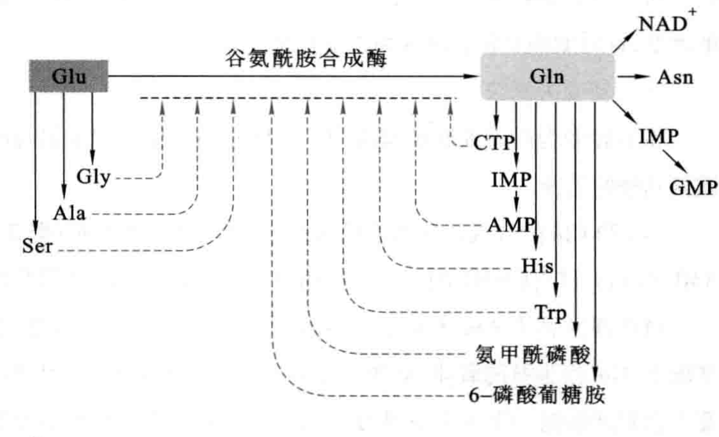


图 31-6　大肠杆菌谷氨酰胺合成酶的活性调节

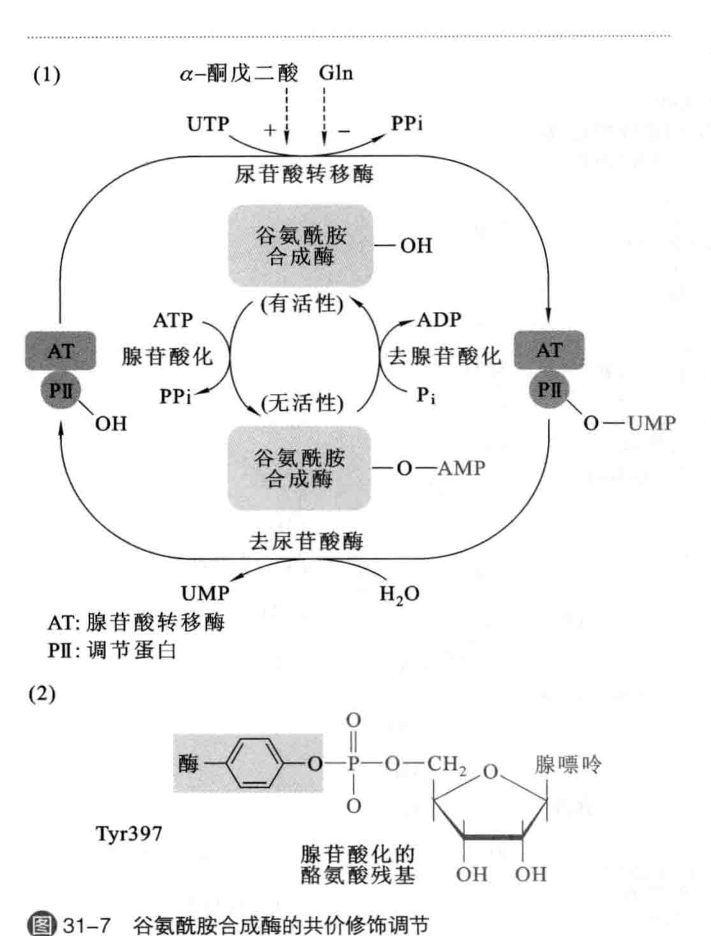


图 31-7 谷氨酰胺合成酶的共价修饰调节

甲酰磷酸、CTP 或 AMP),要么是氨基酸代谢总体状态的指示剂(Ala、Gly 和 Ser)。上述任何一种抑制剂与酶结合都至少能抑制酶的部分活性,结合的抑制剂越多,抑制效果就越强。

大肠杆菌的 GS 还受到共价修饰的调节,但共价修饰的方式并不是真核细胞普遍使用的磷酸化,而是腺苷酸化。如图 31-7 所示,在腺苷酸转移酶(adenylyl transferase,AT)的催化下,GS 上的 Tyr397-OH 接受 ATP 分子上的 AMP 基团,形成无活性的腺苷酸化的 GS。AT 本身的活性又受到调节蛋白 PⅡ的控制,且调节蛋白也可以被共价修饰,其共价修饰的方式为尿苷酸化,修饰酶为尿苷酸转移酶(uridylyl transferase,UT)。α-酮戊二酸刺激 UT 的活性,而 Gln 抑制 UT 的活性。当 UT 被 α-酮戊二酸激活后,与 AT 结合的 PⅡ被尿苷酸化,这时 AT 催化 GS 的去腺苷酸化,GS 恢复活性;相反,在 Gln 的抑制下,UT 缺乏活性,细胞内的去尿苷酸酶(uridylyl-removing enzyme)催化 PⅡ的去尿苷酸化,这时 AT 催化 GS 的腺苷酸化修饰,GS 因此丧失活性。

Quiz3 试设计一个实验,证明大肠杆菌的 GS 发生的是腺苷酸化修饰,而不是磷酸化修饰?

与大肠杆菌的 GS 相比,哺乳动物体内的 GS 也受到调控,但调节的广度远不如微生物,而且没有共价修饰这种方式。哺乳动物 GS 的别构抑制剂有 Gly、Ser、Ala 和氨甲酰磷酸,别构激活剂只有 α-酮戊二酸一种。

(三) 转变为尿素或尿酸

将氨转变成尿素是人类和其他胎盘类哺乳动物解除氨毒的方式。尿素易溶于水,而这些哺乳动物获取水源也不困难,因此溶解在血液中的尿素就很容易通过肾随尿液排出体外。

爬行动物和鸟类则是将氨转变为尿酸,因为尿酸的溶解度不高,容易形成结晶。将氨转变为尿酸对卵生动物是有益处的:一方面,可以使它们能生存在极端干燥的环境之中;另一方面,这对它们的胚胎发育是至关重要的,因为胚胎在羊膜卵内发育的时候,没有水的摄入,形成的尿酸可结晶在尿囊(allantois)内,不需要排出体外就解除了氨毒。

1. 尿素循环

尿素是通过尿素循环(urea cycle)产生的(图 31-8),它发生在哺乳动物的肝细胞。虽然参与尿素循环的个别酶还存在于其他组织,但完整的尿素循环仅存在于肝细胞。肝外组织氨基酸分解产生的氨主要通过 Gln 被转运到肝细胞,但肌肉蛋白分解产生的氨是通过 Ala 循环被 Ala 带往肝细胞。

尿素循环是第一条被发现的循环代谢途径。早在 1932 年,Hans Krebs 和 Kurt Henseleit 发现在有氧条件下,将大鼠肝组织的切片与铵盐保温就可以合成尿素。合成速率可因加入少量的鸟氨酸、瓜氨酸或精氨酸而大大加速,而这 3 种氨基酸的含量并不减少。实验还表明,大量鸟氨酸与肝组织切片及铵盐保温时有瓜氨酸的合成。当时已经知道在排尿素的陆生动物的肝脏中存在一种酶,该酶的催化可导致精氨酸不可逆地水解产生鸟氨酸和尿素。综合以上实验数据,并同时对鸟氨酸、瓜氨酸和精氨酸的化学结构进行比较分析,Krebs 和 Henseleit 提出了尿素合成的循环机制:先是鸟氨酸与氨及 CO_2

Quiz4 肝细胞中有大量的鸟氨酸和瓜氨酸,但它们从来不会参入到蛋白质分子之中,为什么?

图31-8 尿素循环

结合生成瓜氨酸；然后，瓜氨酸再接受一分子氨形成精氨酸；最后，精氨酸水解产生尿素，同时再生出鸟氨酸进入下一轮循环。由于鸟氨酸在循环中的重要性，尿素循环又被称为鸟氨酸循环。

尿素循环中的一部分反应发生在肝细胞的细胞质基质，另一部分反应发生在肝细胞的线粒体基质，现分步加以讨论。

(1) 氨甲酰磷酸的形成　严格地说，这一步反应并不属于尿素循环中的反应，它充其量只是尿素循环的预备反应。催化此反应的酶为氨甲酰磷酸合成酶-Ⅰ(carbamyl phosphate synthetase Ⅰ，CPS-Ⅰ)。该酶位于线粒体基质，负责将谷氨酸脱氢酶产生的氨活化为一个不稳定的高能磷酸化合物，即氨甲酰磷酸，反应共消耗 2 分子 ATP。

除了 CPS-Ⅰ以外，细胞中还有另外一种催化氨甲酰磷酸合成的酶，即 CPS-Ⅱ。与 CPS-Ⅰ不同的是，CPS-Ⅱ存在于所有组织的细胞质基质之中，参与嘧啶核苷酸的从头合成，它是谷氨酰胺依赖型的，即反应中氨的供体是谷氨酰胺(参看第三十二章“核苷酸代谢”)。

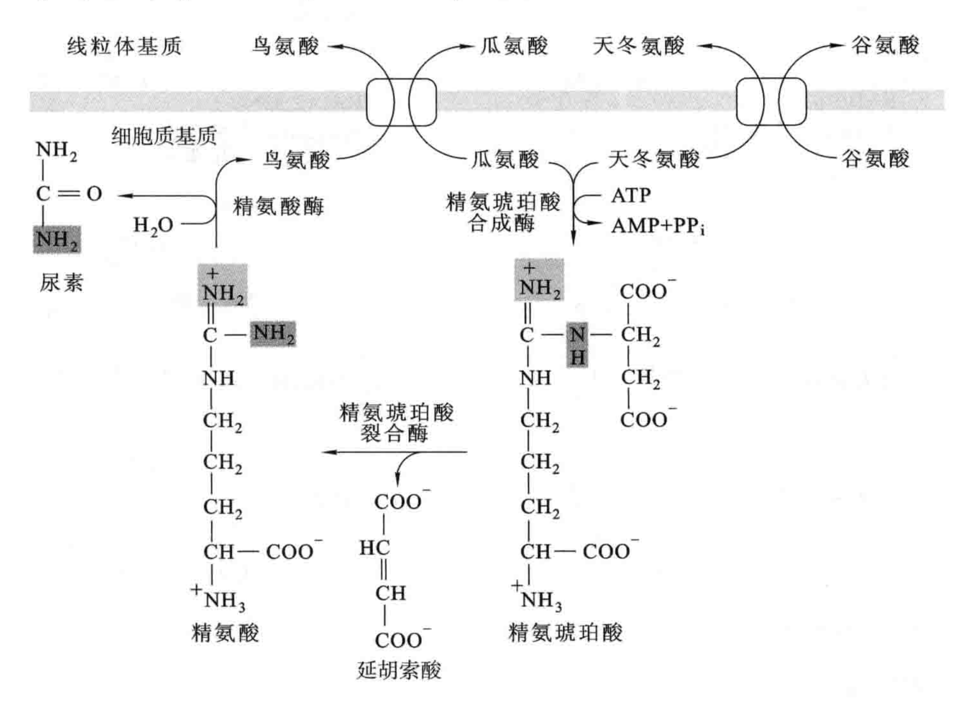
氨甲酰磷酸的合成是肝细胞线粒体的一项主要任务。据估计，CPS-Ⅰ可占肝细胞线粒体基质总蛋白的 20%，与此相同的是，谷氨酸脱氢酶的含量也十分丰富。

(2) 鸟氨酸→瓜氨酸　这也是一步发生在肝细胞线粒体基质内的反应，由鸟氨酸转氨甲酰酶(ornithine transcarbamylase)催化。

$$\text{鸟氨酸} + \text{氨甲酰磷酸} \xrightarrow{\text{鸟氨酸转氨甲酰酶}} \text{瓜氨酸} + \text{磷酸}$$

(3) 鸟氨酸和瓜氨酸的反向转运　尿素循环余下的反应发生在细胞质基质，这需要不断地将瓜氨酸运出线粒体和将鸟氨酸运回线粒体基质。这种转运系统仅存在于肝细胞线粒体内膜上，在特殊的氨基酸转运蛋白即瓜氨酸－鸟氨酸交换体(citrulline-ornithine exchanger)催化下完成。

(4) 谷氨酸－氢氧根和谷氨酸－天冬氨酸的反向转运　尿素的产生同样需要 Glu 持续不断地进入肝细胞的线粒体基质，以补充谷氨酸脱氢酶的底物。此过程是在线粒体内膜上的谷氨酸－氢氧根和 Glu-Asp 的交换体(glutamate-aspartate exchanger)催化下完成的。



(5) 精氨琥珀酸的合成　这一步反应由精氨琥珀酸合成酶(arginino-succinate synthetase)催化。在该酶的催化下，细胞质基质中的瓜氨酸与 Asp 在 ATP 的驱动下缩合成精氨琥珀酸(arginino-succinate)。

$$\text{瓜氨酸} + \text{Asp} + \text{ATP} \xrightarrow{\text{精氨琥珀酸合成酶}} \text{精氨琥珀酸} + \text{AMP} + \text{PPi}$$

Asp 通过这种方式为将来的尿素分子提供了第二个 N 原子。

这是一步高度耗能的反应，在反应中，ATP 被裂解成 AMP 和 PPi，PPi 在焦磷酸酶的催化下，迅速被水解成无机磷酸，因此总反应实际上消耗了 2 分子的 ATP。

(6) 精氨琥珀酸的裂解　这一步反应是由精氨琥珀酸裂合酶(arginino-succinate lyase)催化。在该酶的催化下，精氨琥珀酸裂解成精氨酸和延胡索酸。

$$\text{精氨琥珀酸} \xrightarrow{\text{精氨琥珀酸裂合酶}} \text{精氨酸} + \text{延胡索酸}$$

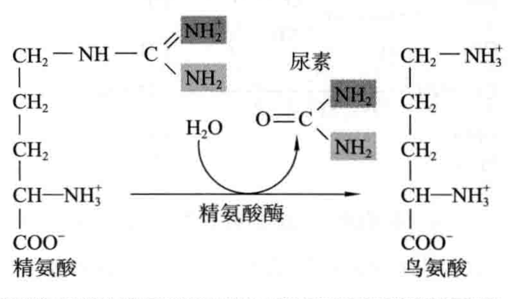
这一步反应与嘌呤核苷酸从头合成途径中的 IMP 转变为 AMP 的反应极为相

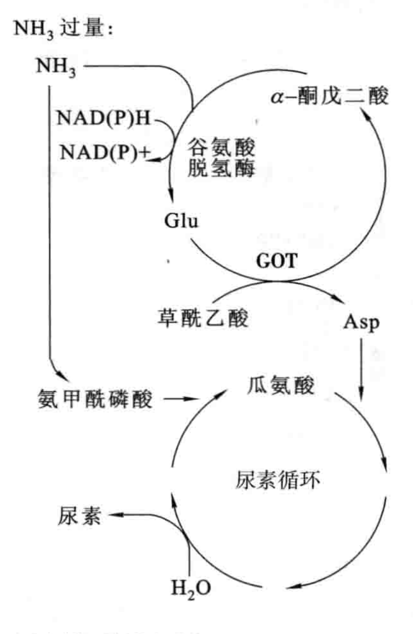
似(参看第三十二章“核苷酸代谢”)。在这两种情况下,延胡索酸均是副产物。生成的延胡索酸在细胞质基质的延胡索酸酶催化下,转变为苹果酸。苹果酸可经草酰乙酸重新转变为 Asp。

(7) 尿素的形成和鸟氨酸的再生　这是尿素循环的最后一步反应。在精氨酸酶(argininase)的催化下,精氨酸被水解成尿素和鸟氨酸,得以再生的鸟氨酸可以进入下一轮循环。

2. 尿素循环的辅助反应

已知尿素分子中的一个 N 原子来自于 Glu 的氧化脱氨或 Gln 水解释放出来的 NH_3,另一个 N 原子来自于 Asp 中的氨基。然而,肝细胞中的 NH_3 和 Asp 的量不一定正好相等,当两者的浓度相差悬殊的时候,肝细胞可以通过图 31–9 所示的辅助反应进行调节,以实现两者量的平衡。如果 NH_3 过量,GOT 催化的反应就有利于 Asp 的形成,以提高 Asp 的量;反之,如果 Asp 的量过量,GPT 催化的反应将有利于 Glu 的形成,以降低 Asp 的量。





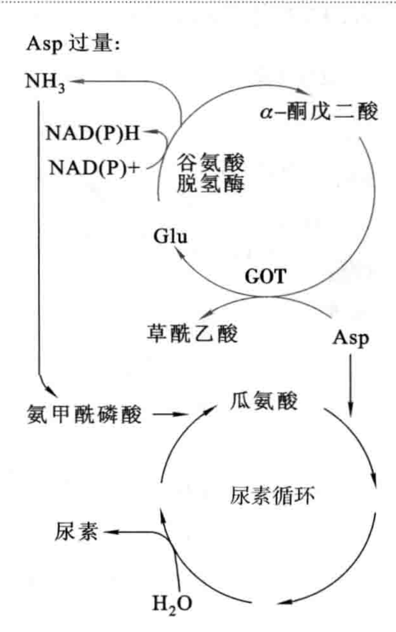


图 31–9　尿素循环的辅助反应

3. 尿素循环的调节

尿素循环的调节有两种方式。

(1) 别构调节　CPS–1 的活性可被 *N*– 乙酰谷氨酸(*N*–acetyl glutamate)强烈激活。事实上,这种激活控制了尿素产生的总速率。目前,人们对 CPS–1 这种奇特的调节方法尚不完全了解,已发现 *N*– 乙酰谷氨酸是细菌鸟氨酸合成的一种中间物,但是这种特性在哺乳动物体内已经丧失,只是调节机制沿袭下来。

在机体内不仅发现了合成 *N*– 乙酰谷氨酸的合成酶,还发现了水解 *N*– 乙酰谷氨酸的水解酶。由这两种酶构成的“无效循环”显然对氮代谢的控制来说十分重要,但其中的机理并不清楚。

(2) 参与尿素循环各种酶浓度的变化　参与尿素循环的各种酶的合成是可以诱导的,底物浓度的升高可刺激它们的合成。

4. 尿素循环的总结

尿素循环的总反应式为:

$$NH_3+HCO_3^-+Asp+3ATP \longrightarrow 尿素+延胡索酸+2ADP+2Pi+AMP+PPi$$

Quiz5 尿素循环和三羧酸循环共同的代谢物是什么?

从总反应式中不难看出,每合成 1 分子尿素,需要消耗 4 分子 ATP。

尿素循环的主要功能是解除氨毒。当肝细胞受损或者尿素循环途径中的任何一种酶有缺陷的时候,尿素合成将会受阻,这将导致血氨升高,引发高血氨症(hyperammonemia)。氨进入脑组织以后,会与脑细

框 31-1 理论联系实际——"冬眠"的生化奥秘

想象一下，外面的温度长时间保持在 -31℃以下，而你一直在睡觉，连续几个月不起床、不吃东西、也不排便。如果你是女性，可能正好处在生育和哺乳期间，可你的后代等你醒来的时候，立即活蹦乱跳。你一定认为，这怎么可能呢？事实上，没有一个人能够做到这一点，但对像北美黑熊或北极熊这样的动物来说，这一切不过是一个正常的冬季下发生的不能再平常的事情了——冬眠。

毫无疑问，在休眠期间，北美黑熊和北极熊使用的主要能源来自其储存的脂肪。由于储存的糖原很少，也许它们的大脑主要靠酮体供能，只需要一点点葡萄糖。然而，有一个代谢过程必须做出适应性的改变，这就是氮代谢。由于氮代谢必然产生尿素，如何将尿素及时排出体外是一个很重要的问题。这两种动物在冬眠期间，会不会通过减少氨基酸的分解代谢来降低尿素的产出呢？答案是否定的。事实上，在冬眠期间，它们体内的氨基酸代谢率保持不变，甚至还有所增加，而血尿素水平却并不高。那么产生的尿素去了哪里？原来尿素在进入膀胱以后，被重吸收到血流之中，然后在通过大肠的时候，被里面的微生物产生的脲酶水解成氨，而氨再被这些微生物用来合成氨基酸。此外，有研究表明，在尿素水解的时候，会在肠道细菌质膜的内外建立一种质子梯度，肠道微生物可用它驱动 ATP 的合成。由此看来，这些肠道微生物对于北美黑熊和北极熊的氮代谢十分重要，正因为如此，一般的情况下这两种动物冬眠得并不深，它们的体温仍远高于环境，结肠内的细菌仍保持活跃的代谢状态。

胞中的 α- 酮戊二酸合成 Glu，并进一步形成 Gln。这必然导致脑细胞中 α- 酮戊二酸浓度的降低，从而削弱了 TCA 循环的能力，也就减少了 ATP 的产率，这可能就是氨中毒可导致昏迷甚至死亡的原因。

三、碳骨架的代谢

碳骨架的代谢有以下几种方式（图 31-10）。

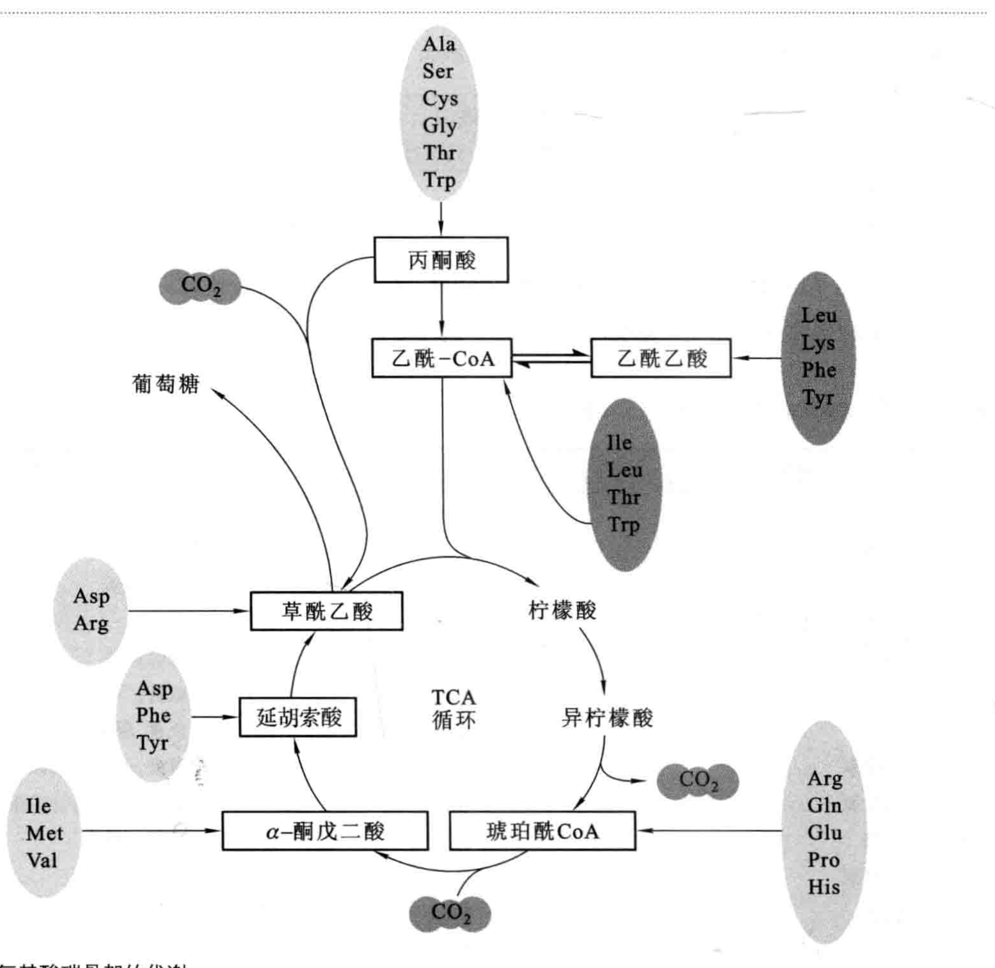


图 31-10 氨基酸碳骨架的代谢

（1）严格生糖　在体内，某些氨基酸的碳骨架被代谢成丙酮酸，或者 TCA 循环途径中的中间物，它们作为糖异生的前体可转变为葡萄糖，这些氨基酸被称为生糖氨基酸（glucogenic amino acid）。20 种常见的蛋白质氨基酸中属于生糖氨基酸的有：Ala、Arg、Asp、Asn、Cys、Gly、Glu、Gln、His、Met、Ser 和 Pro。

（2）严格生酮　Leu 和 Lys 的碳骨架在体内只能被代谢成乙酰 -CoA 或乙酰乙酸，其中前者为酮体和脂肪酸合成的前体，后者本身就是酮体，这两种氨基酸被称为生酮氨基酸（ketogenic amino acids）

Quiz6 如何确定一种氨基酸是生糖还是生酮氨基酸？

（3）生糖兼生酮　Trp、Thr、Tyr、Ile 和 Phe 这 5 种氨基酸的碳骨架较大，其中有一部分在体内可转变为葡萄糖，另一部分可被代谢成乙酰 -CoA 或乙酰乙酸，这些氨基酸属于生糖兼生酮氨基酸。

（4）经 TCA 循环彻底氧化成 CO_2 和 H_2O。

（5）碳骨架可循环利用，经转氨反应重新合成氨基酸。

第二节　氨基酸及其衍生物的合成

Quiz7 苯丙酮酸尿症患者多一种必需氨基酸，你认为这种氨基酸是什么？

不同生物合成氨基酸的能力是不一样的。在有合适的 N 源时，植物和微生物能够从头合成所有的 20 种常见的蛋白质氨基酸，哺乳动物只能合成其中的 10 种非必需氨基酸，部分合成 2 种半必需氨基酸，其余 8 种氨基酸不能被合成，必须从食物中获取。

实际上，任何氨基酸合成的前体都来自于糖酵解、TCA 循环或磷酸戊糖途径，其中 N 原子通过 Glu 或 Gln 进入相关的合成途径，而合成的场所有的在细胞质基质，有的在线粒体。

e31-1 五大家族氨基酸的合成

按照各氨基酸合成前体的性质，除了 His 以外，所有的氨基酸可分为 5 大家族（表 31-1）。这 5 大家族氨基酸合成的基本反应在不同生物体内是相当保守的，而调控机制也主要是通过终产物对限速酶的反馈抑制实现的。

► 表 31-1　五大氨基酸家族（* 表示的不同生物体有所不同）

α- 酮戊二酸家族	丙酮酸家族	甘油酸 -3- 磷酸家族	Asp 家族	PEP 和赤藓糖家族	其他（需要 PRPP）
Glu、Gln、Pro、Arg、Lys*	Ala、Val、Leu	Ser、Gly、Cys	Asp、Asn、Met、Thr、Ile、Lys*	Phe、Tyr、Trp	His

e31-2 氨基酸衍生物的合成

当一种氨基酸被合成后，一方面可以作为寡肽、多肽和蛋白质合成的原料，另一方面还可以作为生物体内多种生物活性物质的前体(表 31-2)。当这些活性物质不能正常合成的时候，可导致机体病变。

► 表 31-2　氨基酸的衍生物及其生物功能

衍生物	氨基酸前体	主要功能
肉碱	Lys	脂肪酸跨线粒体内膜的转运
肌酸磷酸	Gly，Arg	肌肉细胞内的能量贮存
多巴胺	Tyr	神经递质
肾上腺素和去甲肾上腺素	Tyr	激素
GABA	Glu	神经递质
组胺	His	血管扩张
黑色素	Tyr	色素，保护细胞抗紫外
褪黑激素	Trp	激素，促进睡眠和生物节律
一氧化氮	Arg	信号转导，血管扩张
羟色胺	Trp	血管收缩
甲状腺素	Tyr	激素
生长素	Trp	植物激素
乙烯	Met	植物激素
卟啉	Gly，Glu	血红素的前体

推荐网址：

1. http://en.wikipedia.org/wiki/Urea_cycle（维基百科有关尿素循环的内容）
2. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb2/part1/amino.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关氨基酸代谢的内容）
3. http://themedicalbiochemistrypage.org/amino-acid-metabolism.html（完全免费的医学生物化学课程网站，有关氨基酸代谢的内容）
4. http://seqcore.brcf.med.umich.edu/mcb500/aasyl/aametab.html（美国密歇根大学医学院提供的有关氨基酸代谢的内容）
5. http://web.virginia.edu/Heidi/chapter26/chp26frameset.htm（美国弗吉尼亚大学生物化学网络课程，有关氨基酸代谢的内容）

参考文献：

1. Leonard Guarente and et al.SIRT4 Inhibits Glutamate Dehydrogenase and Opposes the Effects of Calorie Restriction in Pancreatic β Cells. *Cell*, 126(5): 941-954, 2006.
2. Kirsch J F, Eliot A C. Pyridoxal phosphate enzymes: mechanistic, structural and evolutionary considerations. *Annu Rev Biochem*, 73, 2004
3. Morris S M. Regulation of enzymes of the urea cycle and arginine metabolism. *Annu Rev Nutr*, 22: 87-105, 2002.
4. Holmes F L. Hans Krebs and the discovery of the ornithine cycle. *Fed Proc*, 39: 216-225, 1980.

第三十二章 核苷酸代谢

核苷酸是一类极为重要的生物小分子，它除了作为核酸的基本组成单位以外，还有一些特殊的功能。例如，ATP 为细胞通用的能量货币，GTP 参与依赖于 G 蛋白的信号转导。已发现人类很多遗传性疾病与核苷酸代谢异常有关，而许多碱基或核苷类似物以及其他抗核苷酸代谢药物，被广泛用于多种疾病的治疗，例如 AZT 治疗艾滋病，因此研究核苷酸代谢以及相关的调控机制具有重要意义。

本章将主要介绍核苷酸的合成代谢、分解代谢、相关的调控机制以及多种抗核苷酸代谢的药物在医学中的应用。

第一节 核苷酸的合成

几乎所有的生物都能自身合成核苷酸，而合成的手段包括从头合成和补救合成两条途径。其中，从头合成是指从最简单的小分子，如 CO_2 和氨基酸开始，经过多步反应，消耗很多的能量，最后生成核苷酸的过程；补救途径是指核苷酸部分降解的产物（包括核苷和碱基）被循环利用，重新转变成核苷酸的过程。显然，补救途径涉及的反应和能耗较少。

一、核苷酸的从头合成

（一）嘌呤核苷酸的从头合成

尿酸作为鸟类嘌呤核苷酸分解代谢的终产物，还保留着嘌呤环的核心结构，其难溶于水、易于分离的性质使它非常适合用于研究嘌呤核苷酸的从头合成。可以使用放射性同位素，对参与嘌呤环合成的小分子前体进行标记，以确定各个小分子前体对尿酸分子结构的贡献，在此基础上可最终确定嘌呤环上各个原子的来源。

John Buchanan 和 Robert Greenberg 就各自使用放射性同位素标记的小分子营养物喂养鸽子，通过分析鸽子粪便中尿酸分子上的同位素分布，率先确定了嘌呤环上各原子的来源（图 32-1）。

e32-1 嘌呤环从头合成各原子来源记忆秘诀

图 32-1 嘌呤环上各原子的来源

嘌呤核苷酸从头合成的小分子前体包括：CO_2、Asp、Gly、Gln 和 N^{10}- 甲酰四氢叶酸。整个合成途径由十多步反应构成，有些反应消耗 ATP 或 GTP。现分别加以叙述。

1. 5- 磷酸核糖 -1-α- 焦磷酸（5-phosphoribosyl-1-α-pyrophosphate，PRPP）的形成

这是一步核糖被活化的反应，由磷酸核糖焦磷酸激酶（ribose-5-phosphate pyrophosphokinase）催化，此酶也称为 PRPP 合成酶（PRPP synthetase），反应的产物为 PRPP。作为主要底物的 5- 磷酸核糖有两种来源，一是来自磷酸戊糖途径，二是来自核苷酸的降解产物——1- 磷酸核糖的异构化。

PRPP 是一种极为重要的代谢中间物，它不仅参与嘌呤核苷酸和嘧啶核苷酸的从头合成，还参与它们的补救合成，此外，它还参与某些核苷酸类辅酶（辅酶Ⅰ和Ⅱ）和某些氨基酸（His 和 Trp）的合成。因此，细胞内的 PRPP 的浓度受到严格的控制，其浓度通常较低。能够抑制 PRPP 合成酶活性的物质有：ADP、2，3- 二磷酸甘油酸、AMP、GMP 和 IMP。

磷酸核糖焦磷酸激酶 (Mg^{2+})ATP → AMP

2. 5-磷酸-1-β-核糖胺(5-phospho-1-β-ribosylamine)的形成

这是一步由谷氨酰胺：PRPP氨基转移酶(Gln：PRPP amidotransferase)催化的反应。在反应中，PRPP中的焦磷酸基团被Gln侧链上的氨基取代。与此同时，核糖C1的构型从α变成天然核苷酸特有的β型。

$PRPP+H_2O$ —Gln:PRPP氨基转移酶, Gln → Glu, (Mg^{2+})→ 5-磷酸-1-β-核糖胺 +PPi

作为嘌呤核苷酸从头合成的限速步骤,该反应受到高度的调控,调节的方式是有活性的单体和无活性的二聚体的互变。核苷酸合成的中间产物IMP和终产物(GMP和AMP)促进单体向二聚体的转变,而反馈抑制此步反应,反应的底物PRPP则激活该酶的活性。

3. 嘌呤核苷酸的形成

在5-磷酸-1-β-核糖胺生成以后,紧接着将再发生9步反应才能形成嘌呤核苷酸。但第一种被合成的嘌呤核苷酸却是次黄苷酸(IMP)。AMP和GMP均由IMP修饰而成。

e32-2 IMP的合成

在从IMP转变成AMP的过程中,首先是IMP与Asp在腺苷酸琥珀酸合成酶(adenylosuccinate synthetase)催化下缩合成腺苷酸琥珀酸,同时消耗1分子GTP;随后在腺苷酸琥珀酸裂合酶(adenylosuccinate lyase)催化下,腺苷酸琥珀酸裂解成AMP和延胡索酸。从IMP转变成GMP的过程中,则首先是在次黄苷酸脱氢酶的催化下,IMP被氧化成黄苷酸(XMP);随后,在鸟苷酸合成酶的催化下,XMP接受Gln侧链上的氨基(原核生物为游离的氨),形成GMP,同时1分子ATP被水解成AMP和PPi(图32-2)。

Quiz1 这步反应与哪一条代谢途径中的哪一步反应相似?

值得注意的是:无论是IMP转变为AMP,还是IMP转变为GMP都是消耗高能磷酸化合物的。但是,在IMP转变为AMP的反应中,消耗的是GTP,而在IMP转变为GMP的反应中,消耗的是ATP。交叉使用核苷三磷酸的好处是可以有效地调节细胞内[GTP]和[ATP]之间的比例,避免两者浓度的失调。

Quiz2 细胞内次黄苷酸转变成鸟苷酸和腺苷酸为什么要放到核苷单磷酸的水平,而不放到核苷三磷酸的水平?

(二)嘧啶核苷酸的从头合成

嘧啶核苷酸从头合成的前体包括Gln、CO_2和Asp等。它与嘌呤核苷酸从头合成至少有两个显著的差别:①嘌呤核苷酸是先形成β-N-糖苷键,然后再逐步形成嘌呤环。而嘧啶核苷酸是先形成嘧啶环,然后再与PRPP形成β-N-糖苷键;②嘌呤核苷酸从头合成的所有反应都是在细胞质基质内发生的,而嘧啶核苷酸从头合成的有些反应则发生在线粒体。

以UMP为例,嘧啶核苷酸的从头合成共有6步反应(图32-3),现分述如下:

1. 氨甲酰磷酸的合成

这是一步由氨甲酰磷酸合成酶Ⅱ(CPS-Ⅱ)催化的反应,是哺乳动物嘧啶核苷酸从头合成途径的限速步骤,类似的反应已在尿素循环中见过。与尿素循环中的氨甲酰磷酸的合成反应不一样的是:①反应场所不同。后者仅发生在肝细胞的线粒体基质,由氨甲酰磷酸合成酶Ⅰ催化(CPS-Ⅰ),前者发生在大多数细胞的细胞质基质;②CPS的调节机制不同。CPS-Ⅰ受到N-乙酰谷氨酸的激活,而CPS-Ⅱ不受N-乙酰谷氨酸的影响,但受到UTP的反馈抑制;③氨基供体不同。CPS-Ⅰ的氨基供体主要来自Glu的氧化脱氨,而CPS-Ⅱ的氨基供体来自Gln的酰胺基。

Quiz3 嘌呤核苷酸和嘧啶核苷酸从头合成的共同前体有哪些?

2. 氨甲酰天冬氨酸的形成

Asp+GTP
IMP
NAD$^+$
GDP+Pi
H_2O
NADH+H$^+$
腺苷酸琥珀酸合成酶
次黄苷酸脱氢酶
腺苷酸琥珀酸
黄苷酸(XMP)
H_2O
Gln+ATP
乌苷酸合成酶
Glu+AMP+PPi
腺苷酸琥珀酸裂合酶
延胡索酸
AMP
GMP

图32-2 从 IMP 合成 AMP 和 GMP

这一步反应由天冬氨酸转氨甲酰基酶(ATC)催化,产物为氨甲酰天冬氨酸,细菌中此步反应为限速步骤。

3. 二氢乳清酸(dihydroorotate)的形成

这是一步分子内的缩合反应,由二氢乳清酸酶(dihydroorotase)催化,氨甲酰天冬氨酸经脱水闭环形成二氢乳清酸。

4. 乳清酸(orotate)的形成

这一步反应由二氢乳清酸脱氢酶催化。在真核细胞中,该酶位于线粒体内膜。二氢乳清酸经脱氢反应转变为乳清酸,NAD$^+$ 为电子受体。

5. 乳清苷酸(orotidylate,OMP)的形成

这一步反应与嘌呤核苷酸的补救途径相似,在乳清酸磷酸核糖转移酶(orotate phosphoribosyltransferase)的催化下,乳清酸与 PRPP 反应形成乳清苷酸。

6. UMP 的形成

Quiz4 在嘌呤核苷酸和嘧啶核苷酸从头合成中,最先合成的核苷酸各是哪一种?

这是一步脱羧反应,由乳清苷酸脱羧酶催化,产物为 UMP。

现已明确,哺乳动物细胞催化前三步反应的酶,即 CPS-Ⅱ、ATC 和二氢乳清酸酶位于同一条多肽链上,是一种多功能酶。与此相类似的是,催化反应 5 和反应 6 的酶,即乳清酸磷酸核糖转移酶和

图32-3 UMP 的从头合成

OMP 脱羧酶也位于同一条多肽链上。

胞苷酸由尿苷酸衍生来，在转变之前，UMP 需要转变成 UTP，后者在 CTP 合成酶的催化下，转变成 CTP，反应式为：UTP+Gln+ATP ⟶ CTP+Glu+ADP+Pi。

Quiz5 试总结一下，到现在为止你接触过多少种多功能酶？

二、核苷酸的补救合成

核苷酸在降解过程中不断产生碱基或核苷，细胞有时会利用现成的碱基或核苷重新合成核苷酸不仅能够节约时间，还能节省能量，因此，补救途径的重要性不亚于从头合成。有证据表明，脑细胞主要通过补救途径合成嘌呤核苷酸。另外，据估计约 90% 的嘌呤碱基可用于补救合成。

（一）嘌呤核苷酸的补救合成

有两类酶可参与嘌呤核苷酸的补救合成：一类属于嘌呤碱基磷酸核糖转移酶，需要嘌呤碱基和 PRPP。它们包括腺嘌呤磷酸核糖转移酶（adenine phosphoribosyltransferase，APRT）和次黄嘌呤 - 鸟嘌呤磷酸核糖转移酶（hypoxanthine-guanine phosphoribosyltransferase，HGPRT），这两种酶分别用于 AMP 和 IMP/GMP 的补救合成。另一类包括核苷磷酸化酶和核苷激酶，其中核苷磷酸化酶可将碱基和 1-磷酸核糖转变为核苷，核苷激酶再将核苷直接激活为核苷酸。

$$\text{腺嘌呤}+\text{PRPP}\xrightarrow{\text{APRT}}\text{AMP}+\text{PPi}$$

$$\text{鸟嘌呤/次黄嘌呤}+\text{PRPP}\xrightarrow{\text{HGPRT}}\text{GMP/IMP}+\text{PPi}$$

（二）嘧啶核苷酸的补救合成

参与嘧啶核苷酸补救合成的酶主要有核苷磷酸化酶和核苷激酶（如胸苷激酶）。许多核苷类药物都是通过此途径在细胞内转变为相应的核苷酸的。

此外，类似于 APRT 或 HGPRT 的嘧啶磷酸核糖转移酶（pyrimidine phosphoribosyltransferase）也可以催化嘧啶核苷酸的补救合成。

$$\text{碱基}+1\text{-}\text{磷酸核糖}\xrightarrow{\text{核苷磷酸化酶}}\text{核苷}+\text{Pi}$$

$$\text{核苷}+\text{ATP}\xrightarrow{\text{核苷激酶}}\text{核苷酸}+\text{ADP}$$

三、核苷二磷酸（NDP）和核苷三磷酸（NTP）的合成

在核苷单磷酸激酶（nucleoside monophosphate kinases）的催化下，NMP 可被激活为 NDP，不同的 NMP 激酶对碱基的特异性不同，但对戊糖无特异性，即不论是核糖还是脱氧核糖。例如，腺苷酸激酶作用于 AMP 和 dAMP，鸟苷酸激酶则作用于 GMP 和 dGMP。

$$\text{AMP+ATP}\xrightarrow{\text{腺苷磷酸激酶}}\text{2ADP}$$

$$\text{GMP+ATP}\xrightarrow{\text{鸟苷磷酸激酶}}\text{GDP+ADP}$$

而核苷三磷酸的合成则应分别考虑：ATP 可以通过底物水平的磷酸化、氧化磷酸化和光合磷酸化直接生成，而一部分 GTP 也可以在 TCA 循环中通过底物水平磷酸化直接产生，其他 NTP 是在核苷二磷酸激酶（nucleotide diphosphate kinase）的催化下形成的。与核苷单磷酸激酶不同的是，核苷二磷酸激酶对碱基和戊糖（核糖还是脱氧核糖）均无特异性。

$$N_1TP+N_2DP\xrightarrow{\text{核苷二磷酸激酶}}N_1DP+N_2TP$$

在有氧条件下，细胞内的 ATP 浓度远远高于其他几种核苷三磷酸，因此，作为磷酸供体的核苷酸几乎全是 ATP。

四、脱氧核苷酸的合成

细胞是先合成核糖核苷酸，后由核糖核苷酸还原成脱氧核苷酸。这是支持“RNA 世界”假说的一个重要证据。催化脱氧核苷酸合成的酶是核苷酸还原酶（riibonuleotide reductase），已发现有三类：

第一类分布最广，存在于绝大多数生物中，其底物为 NDP（图 32-4），因此又称为 NDP 还原酶。这类酶的反应机制涉及活性中心的一个以自由基形式存在的酪氨酸残基，还需要一个双铁氧桥（diferric oxygen bridge）的帮助，以维持上述自由基的再生。

NADPH +H⁺ → NADP⁺ +H₂O；NDP →（NDP 还原酶）→ dNDP

四种 NDP → 四种 dNDP；CTP → CDP → 四种 NDP；AMP,GMP,UMP → 四种 NDP

图 32-4 脱氧核苷酸的合成

第二类存在于蓝细菌、某些细菌和小眼虫（*Euglena*）等生物中，其底物为 NTP，利用脱氧腺苷钴胺素产生催化性的自由基。

第三类仅发现在兼性和专性厌氧菌中，其底物为 NTP，利用 *S*- 腺苷甲硫氨酸和一个铁硫中心，在一个 Gly 残基上产生催化性的自由基。

以大肠杆菌的 NDP 还原酶为例，在反应中 NDP 被还原成 dNDP，直接的还原剂为硫氧还蛋白(thioredoxin)，而最终的还原剂为 NADPH。其中 NADPH 在硫氧还蛋白还原酶(thioredoxin reductase)的催化下，负责还原性硫氧还蛋白的再生(图 32–5)。

NDP 还原酶的作用机制可分为以下几步(图 32–6)①3′–核苷酸自由基的形成，自由基从 R2 亚单位的 Tyr 残基(Y)转移到核糖的 3′–碳上；②脱水，来自 R1 亚单位的一个 Cys 巯基上的氢与核糖 2′–羟基缩合成水；③3′–核苷酸自由基的再生，另一个 Cys 残基巯基上的氢转移到核糖环上，同时在上述两个 Cys 残基之间形成二硫键；④Tyr 自由基的再生，自由基从核糖的 3′–碳转移到 Tyr 残基上，以形

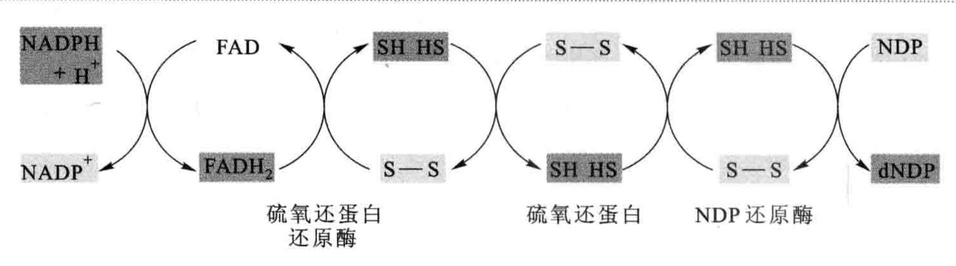


图 32–5 NDP 的还原

图 32–6 NDP 还原酶的作用机理

成稳定的 dNDP；⑤dNDP 被释放，二硫键被硫氧还蛋白或谷氧还蛋白(glutaredoxin)还原成自由的巯基。

NDP 还原酶的活性受到严格的调控，dNDP 的合成通常在真核细胞细胞周期中的 G_1 后期进行。

羟基脲(hydroxyurea)是 NDP 还原酶的特异性抑制剂，它可以有效地抑制 dNDP 的合成，从而达到阻断 DNA 复制的目的。

五、胸苷酸的合成

胸苷酸是在胸苷酸合酶(thymidylate synthase)的催化下，由 dUMP 甲基化而成，甲基供体为 N^5，N^{10}– 亚甲基四氢叶酸(图 32–7)。N^5，N^{10}– 亚甲基四氢叶酸供出亚甲基以后转变为二氢叶酸，二氢叶酸再被 NADPH 还原成四氢叶酸。四氢叶酸则在丝氨酸转羟甲基酶的催化下，与丝氨酸反应重新转变为 N^5，N^{10}– 亚甲基四氢叶酸。由此可见，细胞是先有 RNA 特有的碱基 U，后有 DNA 特有的碱基 T，这可被视为“RNA 世界”假说的又一个重要证据。

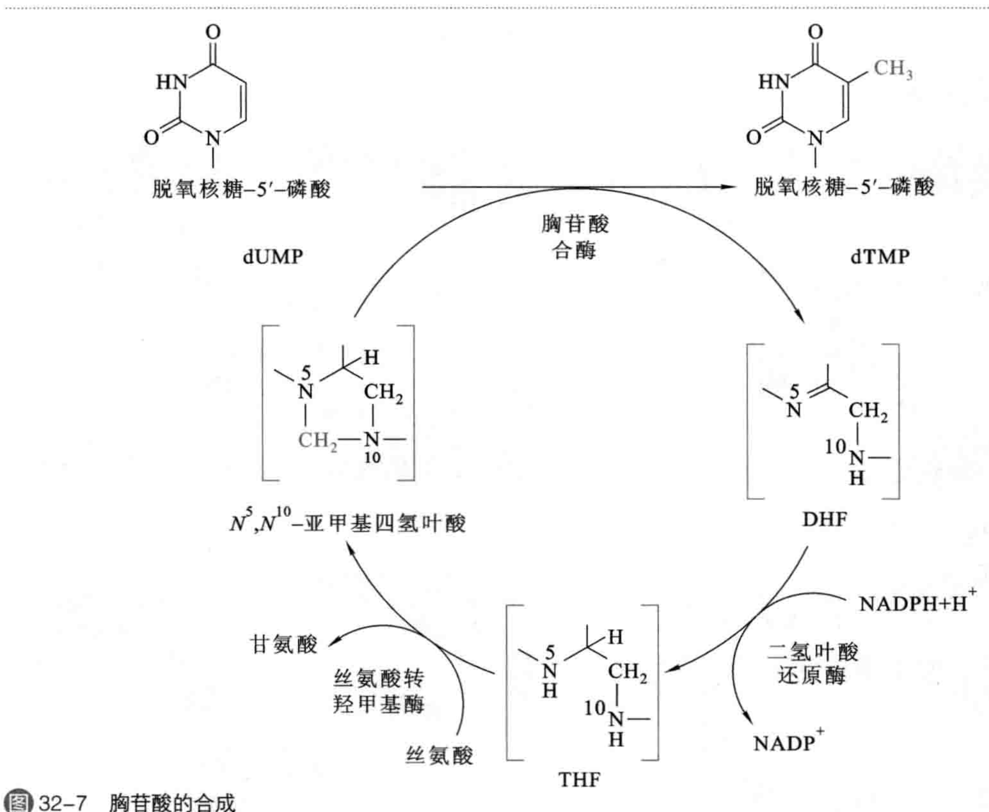


图 32–7　胸苷酸的合成

dUMP 由 dUTP 水解而成，dUTP 酶(dUTPase)催化此反应，反应式为：dUTP → dUMP + PPi。由于反应产生的 PPi 受焦磷酸酶的作用会迅速水解，使反应非常有利于 dUMP 的形成。于是，细胞内的[dUTP]甚低，这就减少了它在 DNA 复制中参入到 DNA 链上的可能性。

细胞生成 dUTP 有两条途径。其一是 UMP → UDP → dUDP → dUTP；其二是由 dCTP 脱氨而成(图 32–4)。

六、核苷酸类辅酶的合成

e32–1 核苷酸类辅酶的生物合成

NAD$^+$、NADP$^+$、FAD 和 CoA 四种辅酶的分子结构之中均含有核苷酸成分，它们在体内的合成与相关维生素的生理功能紧密相连。

框 32–1　生化聚焦——噬菌体对宿主菌内核苷酸代谢途径的改造

在一些噬菌体的基因组 DNA 上，有时可以看到一些修饰的碱基。两种常见的修饰碱基是 U 和 5–羟甲基尿嘧啶(hydroxymethyuracil，hmU)。例如，枯草杆菌的 PBS2 噬菌体基因组 DNA 上完全用 U 代替

了 T，而大肠杆菌 T4 噬菌体基因组 DNA 上含有大量的 hmU 代替 T。这些噬菌体这么做的目的是保护自己的基因组免受宿主细胞编码的限制性内切酶对其进行水解。假如 T4 噬菌体 DNA 上没有这些修饰碱基时，其上的 GAATTC 很容易被大肠杆菌编码的限制性内切酶 *Eco*R Ⅰ识别并切割，但如果 T 换成了 hmU，*Eco*R Ⅰ不能被识别了。

但噬菌体 DNA 上的 U 和 hmU 是如何参入的呢？原来两种噬菌体基因组编码了几种酶，对宿主细胞内原有的核苷酸合成代谢途径进行了改造。改造的目的是降低宿主细胞内 dTTP 的水平，同时提高 dUTP 或 hmdUTP 的水平，这样 dUTP 和 hmdUTP 在噬菌体 DNA 复制的时候，可代替 dTTP 直接参入到 DNA 分子之中。例如，T4 噬菌体编码一种 dTTP 酶，可将 dTTP 水解成 dTMP 和 PPi，而 PPi 很容易被宿主细胞内的焦磷酸酶降解成 Pi。此外，T4 噬菌体还编码一种抑制蛋白，阻止 dUMP 变成 dTMP。再如，PBS2 噬菌体不仅编码一种将 dCTP 脱氨基成 dUTP 的酶，还编码一种催化 dTMP 水解成 TdR 的酶（图 32-8）。

图 32-8　噬菌体对宿主菌核苷酸合成代谢途径的改造

第二节　核苷酸合成的调节

细胞内核苷酸的合成受到严格的调控。调控的主要目的是保持细胞内各种核苷酸之间的浓度平衡，这对于维持 DNA 复制和转录的忠实性十分重要（参看第三十三章“DNA 复制”）。调控的主要手段是终产物的反馈抑制。

一、嘌呤核苷酸合成的调节

嘌呤核苷酸从头合成途径中受到调节的酶有：PRPP 合成酶、谷氨酰胺：PRPP 氨基转移酶、腺苷酸琥珀酸合成酶和次黄苷酸脱氢酶，其中谷氨酰胺：PRPP 氨基转移酶为最重要的限速酶（图 32-9）。

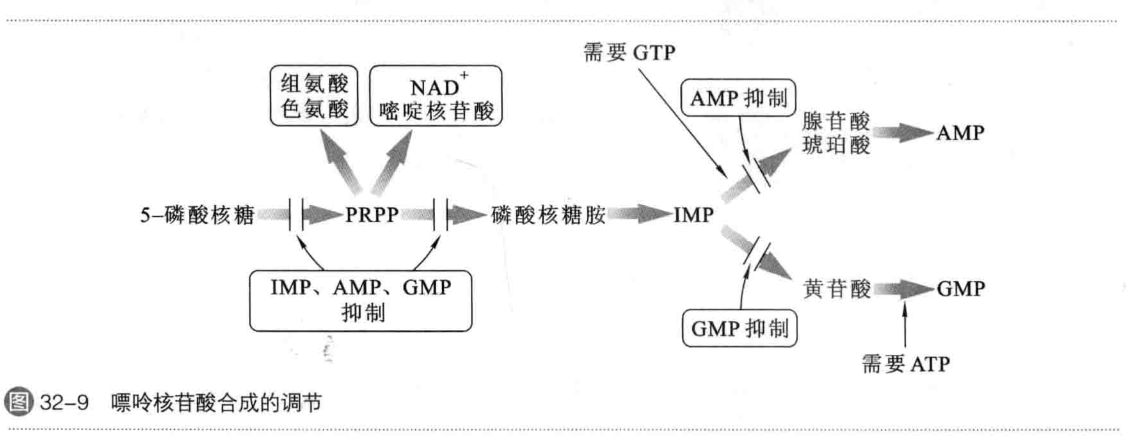


图 32-9　嘌呤核苷酸合成的调节

IMP、AMP 和 GMP 既能反馈抑制 PRPP 合成酶的活性，又能抑制谷氨酰胺：PRPP 氨基转移酶的活性。而作为底物的 PRPP 激活谷氨酰胺：PRPP 氨基转移酶的活性，从而直接启动了嘌呤核苷酸的从头合成途径。

IMP 作为嘌呤核苷酸合成的重要分支点，其后的两条支路分别合成 AMP 和 GMP，而 AMP 和 GMP 分别抑制两条支路的限速酶，即腺苷酸琥珀酸合成酶和次黄苷酸脱氢酶，这又是反馈抑制的一个实例。此外，如前面所述，从 IMP 到 GMP 和 AMP 交叉消耗 ATP 和 GTP 也是一种内在的调节机制。

二、嘧啶核苷酸合成的调节

Quiz6 为什么细菌选择 ATC、哺乳动物选择 CPS-II作为各自合成嘧啶核苷酸的限速酶？

细菌和哺乳动物在嘧啶核苷酸合成的调节位点上并不相同，现分别介绍。

(1) 细菌　细菌嘧啶核苷酸合成的限速酶为 ATC，其中 CTP 和 UTP 为它的反馈抑制剂，ATP 为别构激活剂（参看第十二章“酶活性的调节”）。

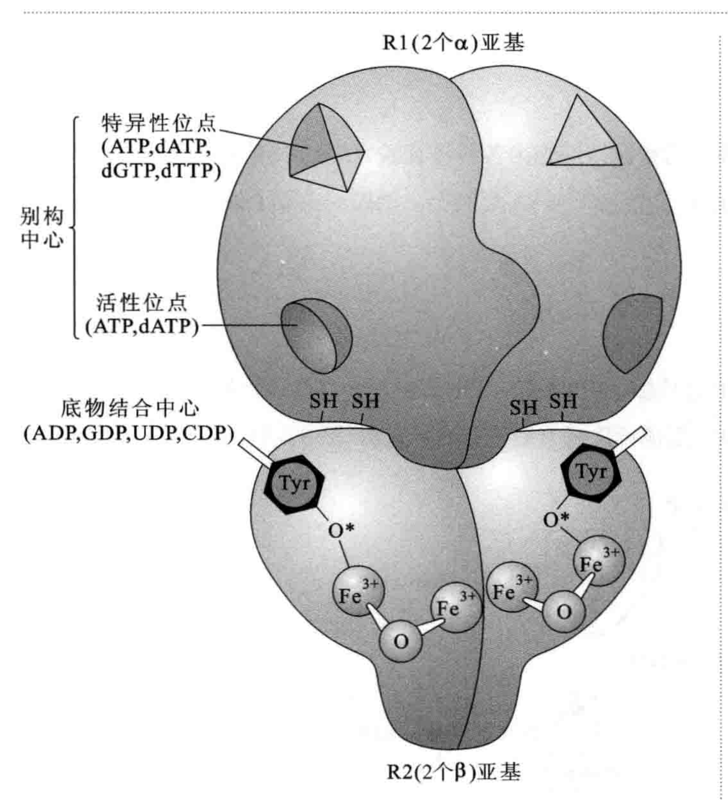
(2) 哺乳动物　CPS-Ⅱ是哺乳动物嘧啶核苷酸合成的限速酶，它一方面受 UDP 或 UTP 的抑制，另一方面受 PRPP 的激活。EGF 能诱导 CPS-Ⅱ的磷酸化，降低其对 UTP 抑制的敏感性，但却也增强了其对 PRPP 激活的敏感性。此外，乳清苷酸脱羧酶也是一个调节位点，其活性受到 UMP 的抑制。

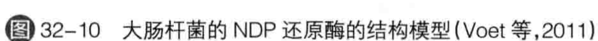
三、脱氧核苷酸合成的调节

脱氧核苷酸合成的调节极为巧妙和精细，其调控位点为核苷酸还原酶，现以大肠杆菌的 NDP 还原酶为例，详细介绍此酶的调控机制。

大肠杆菌的 NDP 还原酶由 R1 和 R2 两个亚单位组成，其中 R1 由 2 个 α 亚基组成，R2 由 2 个 β 亚基组成(图 32–10)。每个 α 亚基含有 1 个特异性位点(specificity site，S 位点)和 1 个活性位点(activity site，A 位点)。α 亚基和 β 亚基之间的缝隙为底物结合位点。

如图 32–11 所示，ATP 和 dATP 与 A 位点的结合分别开启和关闭还原酶的活性，ATP、dATP、dGTP 和 dTTP 与 S 位点的结合则控制哪一种 NDP 与底物结合中心结合。




图 32–10　大肠杆菌的 NDP 还原酶的结构模型（Voet 等，2011）

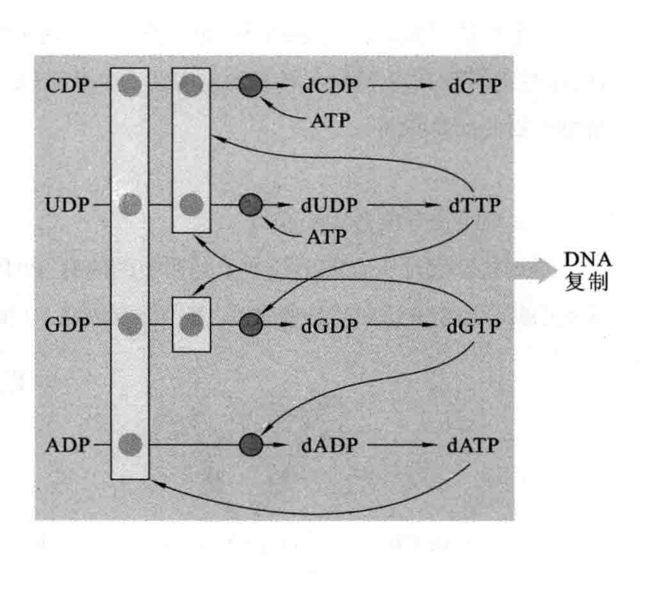


图 32–11　核苷酸还原酶活性的调节机制

在细胞能量状态较高的情况下，ATP 与 A 位点结合以打开还原酶的活性，同时 ATP 与 S 位点的结合促使 CDP 或 UDP 与底物结合中心的结合，致使[dCDP]和[dUDP]升高。dCDP 和 dUDP 通过前面所述的反应转变成 dTTP。当[dTTP]上升到一定程度时，就会占据 S 位点，促使 GDP 与底物结合位

点的结合，即 GDP → dGDP → dGTP。当[dGTP]上升到一定程度时，就会占据S位点，促进ADP与底物结合位点的结合，ADP → dADP → dATP。当[dATP]上升到一定程度时，就取代A位点上的ATP，从而关闭还原酶的活性。

第三节　核苷酸的分解

在核苷酸酶（nucleotidase）、核苷酶（nucleosidase）或核苷磷酸化酶（nucleoside phosphorylase）的依次催化下，核苷酸可分解成碱基和核糖（核糖磷酸），碱基可以进一步分解，核糖则可进入糖代谢。

一、嘌呤核苷酸的分解

图32-12为嘌呤核苷酸在昆虫、爬行类、鸟类和灵长类动物细胞内的降解情况：首先在核苷酸酶的催化下，嘌呤核苷酸脱去磷酸成为嘌呤核苷。随后在嘌呤核苷磷酸化酶的催化下，嘌呤核苷变为嘌呤。嘌呤核苷及嘌呤又可经水解、脱氨及氧化作用生成尿酸。尿酸是这几类生物嘌呤碱基分解的终产物。

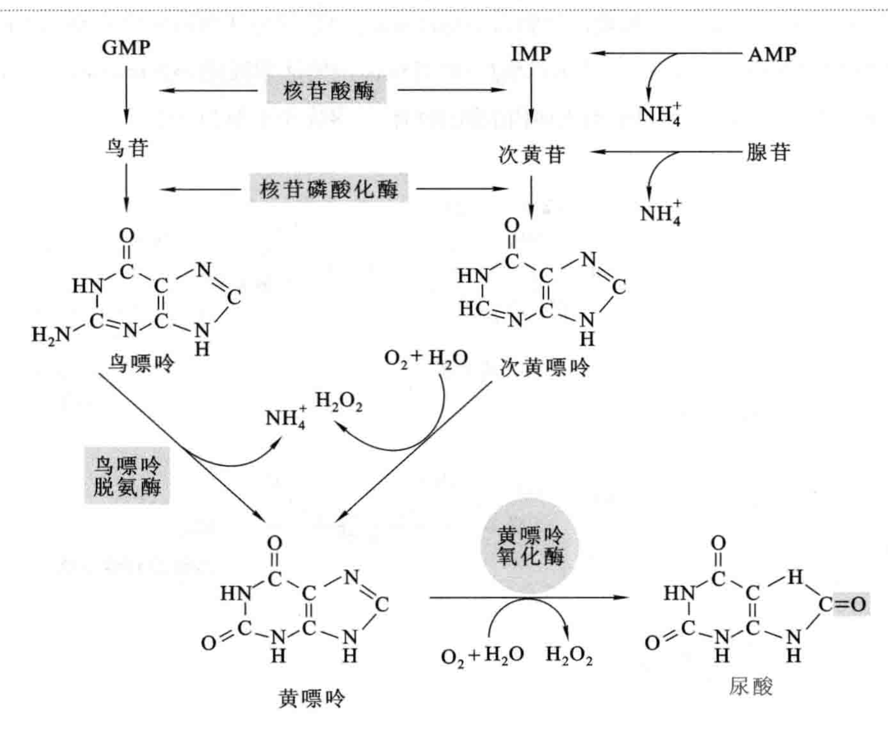


图32-12　嘌呤核苷酸的分解代谢

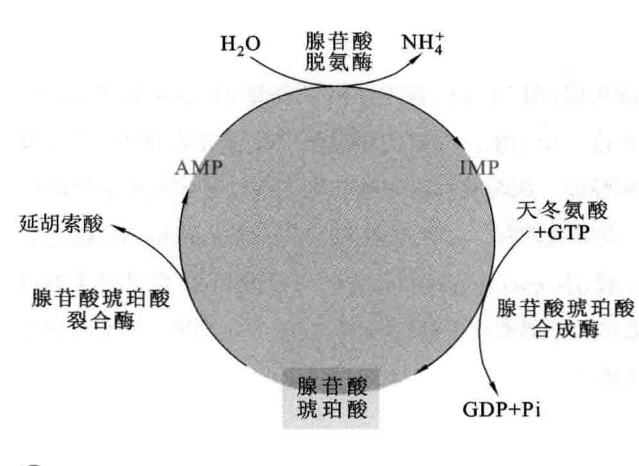


图32-13　嘌呤核苷酸循环

在骨骼肌细胞内，存在一种所谓的嘌呤核苷酸循环（purine nucleotide cycle）（图32-13）：AMP在腺苷酸脱氨酶的催化下转变为IMP，IMP经过腺苷酸琥珀酸再转变成AMP。这种循环能提高延胡索酸的浓度，增强TCA循环，产生更多的ATP以满足肌细胞的需要。

尿酸的形成是在黄嘌呤氧化酶（xanthine oxidase）的催化下，由黄嘌呤直接氧化而成。黄嘌呤氧化酶是一种黄酶，含有FAD、金属钼和铁硫中心，其底

Quiz7 还有哪一种酶也含有钼原子?

物包括次黄嘌呤和黄嘌呤。在反应中,O_2为电子受体,H_2O_2为还原产物,进入尿酸的氧来自水分子(图32-14)。

酶:钼复合物
(Mo^{VI}——钼完全被氧化)
OH^- H^+
酶:钼:底物复合物
(Mo^{IV}——钼完全被还原)
O_2 H_2O_2
$2H^+$
尿酸
+
酶:钼复合物
(Mo^{VI}——钼完全被氧化)

图 32-14 黄嘌呤氧化酶催化反应的机理

但尿酸并非所有动物嘌呤碱基分解的终产物(图 32-15)。大多数哺乳动物在尿酸氧化酶(urate oxidase)的催化下,将尿酸进一步氧化成尿囊素(allantonin),而硬骨鱼体内的尿囊素酶(allantoinase)则将尿囊素水解成尿囊酸(allantoate)。两栖动物和软骨鱼体内的尿囊酸酶(allantoicase)还可将尿囊酸水解成乙醛酸和尿素,而海洋无脊椎动物体内的脲酶则将尿素完全水解为NH_4^+。

(1) 尿酸 —($1/2\ O_2$ $+H_2O$, CO_2; 尿酸氧化酶)→ 尿囊素
灵长类、鸟类、爬行类、昆虫
大多数哺乳动物
(2) 尿囊素 —(H_2O; 尿囊素酶)→ 尿囊酸
硬骨鱼
(3) 尿囊素 —(H_2O, 乙醛酸; 尿囊酸酶)→ 尿素 + 尿素
两栖动物
软骨鱼
(4) 2 尿素 —($2H_2O$, $2CO_2$; 脲酶)→ $4NH_4^+$
海洋无脊椎动物

图 32-15 尿酸的进一步分解

二、嘧啶核苷酸的分解

嘧啶核苷酸的分解代谢途径与嘌呤核苷酸相似(图 32-16):首先在核苷酸酶及核苷磷酸化酶的催化下,依次除去磷酸和核糖,生成的嘧啶碱基再进一步分解。分解代谢过程中有脱氨基、氧化、还原及脱羧基等反应。胞嘧啶经脱氨基反应转变为尿嘧啶。尿嘧啶和胸腺嘧啶则先在二氢嘧啶脱氢酶的催化下,被 NADPH 分别还原为二氢尿嘧啶和二氢胸腺嘧啶。二氢嘧啶酶(dihydropyrimidinase)催化嘧啶环水解,分别生成β-丙氨酸和β-氨基异丁酸(β-aminosiobutyrate)。β-丙氨酸和β-氨基异丁酸可在β-脲基丙酸酶(β-ureidopropionase)、转氨酶和脱氢酶催化下继续分解,最后产生丙酰-CoA 或甲基丙酰-CoA。β-氨基异丁酸亦可随尿直接排出体外。

图 32-16 嘧啶核苷酸的分解

第四节 几种与核苷酸代谢相关的疾病

1. 痛风

痛风(gout)是尿酸过量产生或者尿酸排泄不畅造成的一种疾病，其临床特征为高尿酸血症(hyperuricemia)和反复发作的急性单一关节炎。由于尿酸或尿酸盐的溶解度有限，当其在血浆中的浓度超过某临界值(约 576 μmol/L)的时候，极易形成结晶，并沉积在关节、软组织和肾等处，导致关节炎、尿路结石和肾病等。

参与嘌呤核苷酸代谢的 PRPP 合成酶、谷氨酰胺：PRPP 氨基转移酶或 HGPRT 的缺陷均可以诱发痛风。PRPP 合成酶和谷氨酰胺：PRPP 氨基转移酶是嘌呤核苷酸从头合成的限速酶，这两种酶的任何一种若发生突变，会使它们对反馈抑制不再敏感，引起嘌呤核苷酸过量合成，从而导致尿酸水平异常。

HGPRT 的缺陷与痛风之间的关系尚没有完全确定，但一种可能的解释是：HGPRT 能够消耗 PRPP，降低 PRPP 的浓度。若这种酶有缺陷，一方面嘌呤核苷酸合成的补救途径受阻，而减少了嘌呤碱基的消耗，另一方面 PRPP 浓度的升高可激活其从头合成途径，而增加了嘌呤碱基的合成。这两个方面都增加体内嘌呤碱基的量，从而使尿酸的水平上升。

临床上治疗痛风的特效药物是别嘌呤醇(allopurinol)，它是黄嘌呤氧化酶的一种自杀型抑制剂。当它在细胞内被黄嘌呤氧化酶氧化成别黄嘌呤(alloxanthine)以后，就与酶的活性中心紧密结合从而强烈抑制了它的活性，这样可以有效地抑制了尿酸的产生(图 32-17)。此外，别嘌呤醇还可以与 PRPP

图 32-17　别嘌呤醇治疗痛风的机理

反应生成别嘌呤核苷酸，消耗 PRPP，从而阻止 PRPP 对谷氨酰胺:PRPP 氨基转移酶的激活。

2. 严重联合免疫缺陷病（severe combined immunodeficiency，SCID）

SCID 是腺苷脱氨酶（adenosine deaminase，ADA）单个基因突变引起的。患者的免疫反应几乎完全丧失，必须生活在无菌的环境中，任何病原体的感染都可能是致命的。ADA 的缺陷能导致 SCID 的发生，是因为缺乏 ADA 使细胞内的 dATP 急剧升高。高浓度的 dATP 与核苷酸还原酶的 A 位点结合，关闭该酶的活性，使细胞内 dNDPs 不能有效地合成，这必然影响到细胞内 DNA 的复制（图 32-18）。

图 32-18　ADA 的缺陷对 DNA 复制的影响

受 ADA 突变影响最深的是白细胞，因为任何免疫反应的发生都需要白细胞的分裂，而细胞分裂之前 DNA 首先需要复制，当白细胞因为 DNA 复制受到抑制而不能有效分裂以后，免疫反应也就无从谈起了。

当今 SCID 的治疗除了骨髓移植和基因治疗（用正常的 *ADA* 基因代替有缺陷的 *ADA* 基因）以外，尚没有其他有效的方法。虽然世界上第一次成功的基因治疗就是应用在 SCID 上，但基因治疗费用昂贵，成功率不高，一时还难以推广。

3. Lesch-Nyhan 综合征

Lesch-Nyhan 综合征得名于一名医学院的学生 Michael Lesch 和他的导师 William Nyhan，他们于 1964 年最早报道了这种疾病。Lesch-Nyhan 综合征是一种隐性的性连锁遗传性疾病，此病患者几乎都是男性，女性仅为携带者。该病的病因是由于 HGPRT 有缺陷造成的。Lesch-Nyhan 综合征的主要症状包括：高尿酸血症、肌强直、智力迟钝和自残等。

Quiz8 HGPRT 缺乏为何能够导致高尿酸血症？

4. 乳清酸尿症

乳清酸尿症（orotic aciduria）是一种比较罕见的遗传性疾病，主要症状为尿中有大量乳清酸、重度贫血和生长迟缓。其病因是催化嘧啶核苷酸从头合成的一种双功能酶有缺陷，这种酶同时具有乳清

酸磷酸核糖转移酶和 OMP 脱羧酶的活性。

临床使用尿嘧啶或尿苷治疗，因为尿嘧啶或尿苷在细胞内经补救途径生成 UMP，可抑制 CPS-Ⅱ活性，减少乳清酸的产生。

第五节 常见的抗核苷酸代谢药物

核苷酸作为 DNA 和 RNA 合成的前体，其在细胞内的代谢情况可直接影响到 DNA 的复制和转录，进而影响到细胞的分裂。

抗核苷酸代谢类药物是一些参与核苷酸生物合成的天然代谢物的类似物，它们通常作为竞争性抑制剂或自杀性底物，抑制核苷酸合成的某一种酶，干扰或阻断核苷酸的合成，从而阻断 DNA 的复制和转录，致使细胞的分裂受到抑制。

癌细胞的生长和分裂十分迅速，因此对核苷酸的需求比一般细胞要大得多，它们对核苷酸合成抑制剂的作用就更为敏感，这就奠定了癌症化疗的生化基础。

抗核苷酸代谢类药物除了可用于治疗癌症以外，还经常作为抗病毒的药物。

1. 叶酸类似物

如前面所述，在嘌呤核苷酸的从头合成途径中，由转甲酰酶催化的反应需要 N^{10}- 甲酰四氢叶酸，胸苷酸合成需要 N^5,N^{10}- 亚甲基四氢叶酸，这两种形式的四氢叶酸前体都是二氢叶酸。从二氢叶酸转变为四氢叶酸需要二氢叶酸还原酶，叶酸类似物是二氢叶酸还原酶的抑制剂，如氨基蝶呤（aminopterin）、氨甲蝶呤（methotrexate，MTX）和甲氧苄氨嘧啶（trimethoprim），其中氨基蝶呤和氨甲蝶呤是哺乳动物细胞二氢叶酸还原酶的抑制剂，临床上被用于多种癌症的治疗，而甲氧苄氨嘧啶是细菌二氢叶酸还原酶的抑制剂，是一种抗菌药。

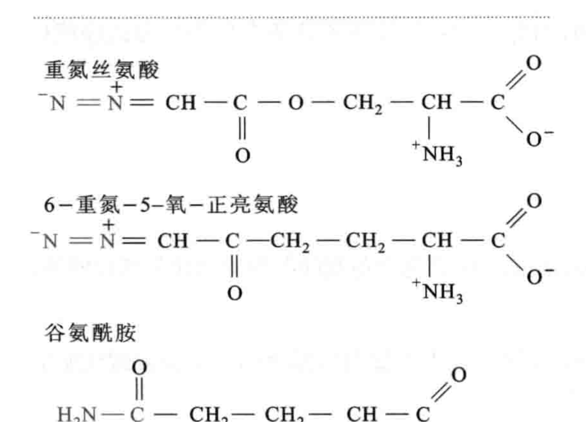


图 32-19 谷氨酰胺的类似物

2. 谷氨酰胺类似物

核苷酸生物合成有好几种酶以 Gln 为底物，包括谷氨酰胺：PRPP 氨基转移酶、鸟苷酸合成酶、CPS-Ⅱ和 CTP 合成酶等。可作为抗生素使用的重氮丝氨酸（azaserine）和 6- 重氮 -5- 氧 - 正亮氨酸（6-diazo-5-oxonorleucine）在结构上是谷氨酰胺的类似物(图 32-19)。它们能够抑制这几种酶的活性，从而阻断细菌核苷酸的合成。

Quiz9 叶酸类似物会影响哪些核苷酸的生物合成?

3. 碱基类似物

（1）嘌呤类似物 有 6- 巯基嘌呤（6-mercaptopurine）、6- 巯基鸟嘌呤（6-mercaptoguanine）和 8- 氮杂鸟嘌呤（azaguanine）等（图 32-20），它们在细胞内经补救途径转变为相应的核苷酸类似物，通过三种方式来抑制嘌呤核苷酸的合成：①抑制 IMP → AMP 或 GMP；②竞争性抑制 HGPRT，阻断嘌呤核苷酸的补救合成；③抑制谷氨酰胺：PRPP 氨基转移酶的活性。其中 6- 巯基嘌呤被长期用于儿童白血病的治疗，治愈率从最初的 4% 左右提高到 80%以上。

（2）嘧啶类似物 主要有 5- 氟尿嘧啶（5-fluorouracil，5-FU）、5- 氟胞嘧啶（5-fluorocytosine）和 5- 氟乳清酸（5-fluroorotate）（图 32-21），但以 5-FU 最常用。

5-FU 在细胞内经补救途径转变为 5- 氟尿苷酸(5-FUMP)。5-FUMP 是胸苷酸合酶的自杀型底物，可不可逆地抑制该酶的活性，从而阻断体内胸苷酸的合成，达到抑癌的效果。

6-巯基嘌呤　　6-巯基鸟嘌呤　　8-氮杂鸟嘌呤

图 32-20　嘌呤类似物

5-氟尿嘧啶　　5-氟胞嘧啶　　5-氟乳清酸

图 32-21　嘧啶类似物

4. 核苷类似物

主要有 3′-叠氮胸苷(AZT)、阿拉伯糖胞苷(cytosine arabinose, araC)和双脱氧肌苷(ddI)(图 32-22),它们在体内经补救途径分别转变为 AZTPP、ddITP 和 araCTP,然后可参入到正在合成的 DNA 链中,抑制链的延伸。其中 AZT 和 ddI 能够有效地阻断 HIV 病毒的逆转录,已成为治疗艾滋病最有效的药物之一,而 araC 主要用于急性白血病的治疗。

AZT　　araC　　ddI

图 32-22　核苷类似物

5. 核苷酸类似物

理论上来说,上述碱基类似物和核苷类似物均可在体外制成相应的核苷酸类似物,但由于核苷酸类似物带有高度的负电荷,难以进入细胞起作用,因此很少直接使用核苷酸类似物作为抗癌或抗病毒的药物。

e32-4 本章小结填填看

推荐网址:

1. http://themedicalbiochemistrypage.org/nucleotide-metabolism.html(完全免费的医学生物化学课程网站,有关核苷酸代谢的内容)
2. http://library.med.utah.edu/NetBiochem/pupyr/pp.htm(美国犹他大学图书馆提供的有关核苷酸代谢的内容)
3. http://oregonstate.edu/dept/biochem/hhmi/hhmiclasses/biochem/lectnoteskga/2kjan31lecturenotes.html(美国俄勒冈州立大学提供的有关核苷酸代谢的内容)
4. http://web.virginia.edu/Heidi/chapter27/chp27frameset.htm(美国弗吉尼亚大学生物化学网络课程,有关核苷酸代谢的内容)

参考文献:

1. Maor Bar-Peled, Malcolm A O'Neil. Plant Nucleotide Sugar Formation, Interconversion, and Salvage by Sugar Recycling. *Annu Rev Plant Biol*, 62:127-155, 2011.
2. Jordan A, Reichard P. Ribonucleotide reductases. *Annu Rev Biochem*, 67:71-98, 1998.
3. Eliasson R, Pontis E, Sun X, Reichard P. Allosteric control of the substrate specificity of the anaerobic ribonucleotide reductase from *Escherichia coli*. *J Biol Chem*, 269:26 052-26 057, 1994.

代谢生物化学思考题

1. 某些生活在严寒地带的植物能够通过特殊的产热机制将雪融化。经研究发现，其体内并没有解偶联蛋白，但构成它的呼吸链的复合体 I 有"替补"，如果使用替补的复合体 I，则发现其 ATP 产率降低。你认为这些植物是如何产热的？你又如何设计一个实验证明你的解释？

2. 当分别将 ^{3}H-NADH 和 ^{14}C-NADH 加到含有线粒体的细胞抽取液中，你认为经过一段时间以后，线粒体基质中会有放射性吗？为什么？

3. 假定你在 pH 7.0 的无营养溶液中，分离得到一种大肠杆菌。这些大肠杆菌细胞因为完全没有营养而不能合成 ATP。如果你突然在溶液中加入盐酸，使溶液的 pH 下降到 4.0，那么细胞内的 ATP 水平会有什么变化？这种变化是暂时的还是持续的？为什么？

4. 预测下列突变对肝细胞糖酵解速率的影响如何。

(1) PFK1 失去 ATP 的别构位点。

(2) PFK1 失去柠檬酸结合位点。

(3) 2,6- 二磷酸果糖双功能酶失去磷酸酶结构域。

(4) 丙酮酸激酶失去 1,6- 二磷酸果糖结合位点。

(5) PFK1 的 ATP 别构位点结合 ATP 的亲和性提高。

5. 一个新来的田径教练上任后，要求他的短跑运动员改吃低糖、高蛋白的食物。结果不久之后运动员成绩开始下降。这个教练员的对于食物的建议有何问题？

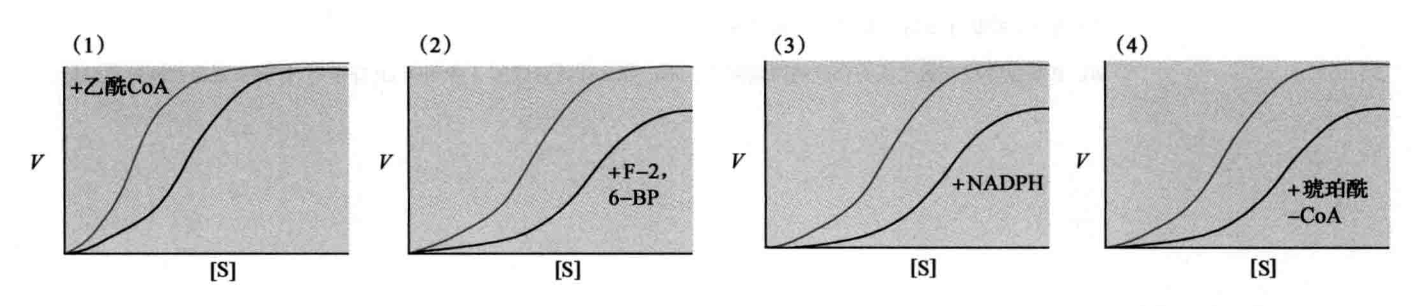
6. 根据柠檬酸合酶的催化机理，你认为当生物进化到某个阶段，是否会出现一种能同时将乙酰 -CoA 分子中高能硫酯键重新转化成 ATP（与琥珀酸硫激酶相似）的柠檬酸合酶？以这种方式恢复能量对三羧酸循环的热力学有何影响？

7. 在大量的丙二酸存在下，将[甲基 -^{14}C]丙酮酸加入到分离的肝细胞中，经过一段时间以后，从肝细胞中分离的异柠檬酸在 C2 和 C5 位置被同位素标记。对此结果，试给予解释。

$$
\begin{array}{l}
{}^{14}CH_2—COO^- \\
\ \ | \\
CH—COO^- \\
\ \ | \\
{}^{14}CH—COO^- \\
\ \ | \\
OH
\end{array}
$$

8. 有许多遗传性代谢缺陷与糖原代谢有关。有一位病人在短暂的禁食后，患低血糖症，他不能进行长时间的运动。病人的肌糖原是正常的，但他的肝糖原比正常人要高许多，尽管它的结构似乎是正常的。你认为哪三种酶的缺乏可能导致上述症状？需要做何种实验加以验证？

9. 以下显示的是几种中间代谢酶催化反应的底物浓度与反应速率的曲线，曲线中的化合物是别构效应物。试根据曲线，写出催化反应的酶、底物和产物。



10. 脂肪酸的氧化是如何被调控的？请解释机体为什么不在脂肪酸的活化或 β 氧化中，选择一步反应作为调节位点？

11. 离体的心肌细胞在缺乏葡萄糖和脂肪酸的情况下，如果提供乙酰乙酸也能正常地收缩。

(1) 乙酰乙酸是如何充当代谢燃料的？

(2) 即使提供大量的乙酰乙酸，其进入三羧酸循环的速率也会逐步下降，除非同时提供一定量的丙酮酸。为什么？

12. 一个学生有只吃奇数碳链脂肪酸的习惯，后来医生发现他缺乏维生素 B_{12}，因此要他停止这种不良习惯，直到维生素 B_{12} 缺乏症治愈以后。请问医生这种治疗方案的生化依据是什么？维生素 B_{12} 缺乏会对脂肪酸代谢造成什么后果？

13. 某些植物的次生代谢物——羟基柠檬酸被用做减肥药物。

```
     CH2—COO-
     |
HO—C—COO-
     |
HO—CH—COO-
```

羟基柠檬酸

(1) 这种化合物能够抑制柠檬酸裂合酶的活性，你认为抑制属于哪一种形式？

(2) 为什么柠檬酸裂合酶活性的抑制能够阻止糖转变成脂肪？

(3) 你认为在体内还有哪些化合物的合成受到羟基柠檬酸的抑制？为什么？

14. 正如大多数其他的氨基酸一样，苏氨酸在体内的分解代谢可用来支持机体在饥饿状态下的生存，而偶数脂肪酸做不到？为什么？奇数脂肪酸有同样的效果吗？

15. 预测下列突变对胆固醇代谢和脂代谢会带来什么影响。

(1) 肉碱：软脂酰转移酶 I 对丙二酸单酰 -CoA 不再敏感。

(2) 将 HMG-CoA 还原酶上磷酸化的位点（一个特殊的 Ser 残基）替换成 Ala。

(3) 过量表达 SREBP 上的碱性螺旋 - 环 - 螺旋结构域（无跨膜螺旋）。

(4) 肝细胞组成型表达 LDL 受体

(5) 使柠檬酸不能与乙酰 CoA 羧化酶结合

16. 有两组实验大鼠，在一个月内分别喂食两种不同的脂肪酸作为唯一的碳源。第一组喂食的是正庚酸，第二组喂食的是正辛酸。在一个月之后，发现两组大鼠有很大的差别：第一组大鼠表现得很健康，体重还有增加；第二组肌肉萎缩无力，体重减轻。试解释导致这两组大鼠出现这些差异的原因。

17. 将脂肪酸合酶能被纯化到均一状态后，先与乙酰 -CoA 和[^{14}C]标记的丙二酸单酰 -CoA 保温，但无 NADPH，然后将过量的[^{14}C]标记的丙二酸单酰 -CoA 透析除去，再添加没有被同位素标记的丙二酸单酰 -CoA 和 NADPH。

(1) 为什么第一步不加 NADPH？

(2) 软脂酸上的哪些 C 被同位素标记？

18. 将纯化的谷氨酸转氨酶透析，可用来去除与酶结合的磷酸吡哆醛(PLP)。但是，PLP 与这种酶的解离非常慢。然而，如果在酶液中加入 Glu，就可提高 PLP 与酶的解离速率，为什么？

19. 在人禁食的最初几天，N 排泄提高，高于正常水平。在几周以后，N 排泄下降到最低水平，而且会继续下降。然而在脂肪储备用完以后，N 排泄又开始上升到较高的水平。

(1) 何种因素诱发 N 排泄的升高？

(2) 为什么几周后 N 排泄开始下降？

(3) 为什么脂肪耗尽后 N 排泄又开始升高？

20. 如果嘌呤核苷酸在从头合成的时候有 ^{15}N-Asp，那么在新合成的 ATP 和 GTP 分子中，什么位置会被同位素标记？

第三篇

分子生物学

分子生物学内容简介及学习方法

从本篇开始,接触到的知识实际上属于分子生物学的内容。我个人认为,本部分的内容是生物化学中逻辑性最强的部分,因此,如果方法得当,可以达到良好的学习效果。要学好这部分的内容,最重要的几点应该是理解、整合、比较和归纳。因为这一部分的内容是围绕"中心法则"展开的,死记的内容大大减少,学习的目的就是要搞清楚"一个中心"(DNA)和"两个基本点"(RNA和蛋白质)之间内在的逻辑关系。在学习中,一定要将它们作为一个整体来看,一环套一环,同时注意科学家们在揭示一些重要发现的时候所使用的各种方法,不时问问自己:如果我自己遇到该怎么办?例如,在学习DNA复制的一般特征的时候,基本上每一要点的背后都隐藏着设计巧妙的实验,从Meselson和Stahl设计的被誉为生命科学中最美丽的实验,即证明Watson和Crick提出的半保留复制的实验,到Okazaki设计的证明DNA复制是半不连续的实验。每一个实验设计得都很巧妙,简直就像创作一件精美绝伦的艺术品!我们在学习的时候,除了要搞清楚这些实验的结果与得出的结论之间的逻辑关系之外,更需要理解这些实验设计的思路,以便为将来自己做科研的时候能从中得到宝贵的借鉴。老师在授课的时候,要多多介绍相关的背景故事,以使同学更容易理解基于这些实验设计而衍生出来的分子生物学真谛。总之,归纳起来有以下几点:

1. 注意将细菌、古菌和真核生物进行比较

无论是细菌、古菌,还是真核生物,它们都要进行DNA复制、转录、转录后加工、翻译等基本的生命过程,这些过程既有相同之处,也有许多差异。在学习的时候,时刻要留意将三大生命系统进行全面的比较,特别需要注意的是,在诸多分子生物学过程和机制中,古菌更接近真核生物。例如:细菌没有组蛋白和核小体结构,但许多古菌与真核生物有组蛋白和核小体结构。再如:在学习弱化子机制的时候,要注意这种机制是原核系统特有的,真核系统没有。那么,为什么真核系统不可能有这种调控系统呢?如果能借鉴此类方法去学习的话,那所有的内容就活了,将它们串在一起理解要比孤立地记忆强得多!

2. 注意将细胞中相关的分子机制进行比较

细胞内的很多分子机制是很相似的,这就需要我们在学习的时候,将相关联的分子机制放在一起去领会、理解。如DNA复制和DNA转录,两者有很多共同的性质:都需要解链,合成的方向都是从5′→3′,都遵循Watson和Crick碱基配对原则。当然,在意识到这些共同性质的时候,也不能忽视它们的差别:DNA复制需要引物,RNA不需要;DNA聚合酶通常具有自我校对能力,RNA聚合酶没有。这里更要明白为什么会有这些差异,为什么允许有这些差异?

3. 以"中心法则"为核心,"碱基互补配对"和"蛋白质与核酸之间的相互作用"为主线,巧妙地利用"外因与内因关系"的理论,全面理解分子生物学的机制

分子生物学的核心内容是所谓的"中心法则",即生物体内的三种生物大分子——DNA、RNA和蛋白质之间的关系。其中涉及遗传信息的复制、损伤修复、重组、转录、逆转录、转录后加工和翻译等。这些过程总是涉及蛋白质和核酸分子之间的相互作用,以及碱基互补配对,因此,掌握蛋白质和核酸分子之间相互作用的规律以及碱基互补配对的原则,对于深入理解分子生物学的各种机制和原理至关重要。另外,细胞内的很多机制都可以使用哲学中"外因"和"内因"关系理论进行理解,掌握这一点非常重要。例如,理解DNA复制为什么具有固定的起点?这涉及DNA复制起始区和复制起始蛋白之间的相互作用,在这里可以将DNA复制起始区看成"内因",复制起始蛋白(大肠杆菌为DnaA蛋白)看成"外因"。按照"内因"和"外因"之间的关系原则,即"内因"是变化的根据,"外因"是变化的条件,"外因"需要通过"内因"起作用,DNA复制区所具有的特殊序列是DNA复制具有固定起点的根本原

因,即“内因”,但仅有它是不够的,还需要识别这种特殊序列的蛋白质,它就是“外因”,正是它们之间的相互作用才使得DNA复制能从固定的起点开始。

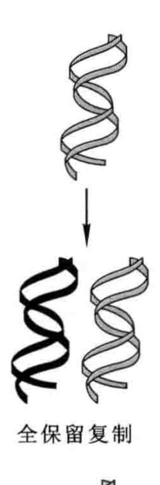
4. 注意掌握各种研究方法的原理及其应用

分子生物学的发展是与研究方法的进步分不开来的,而反过来它的发展又使得人们能提出和发明新的研究手段。两者之间相互依存,相互促进。因此,在学习各章节内容的时候,对于分子生物学家在研究各种分子机理时所使用的方法要充分理解。例如,对参与DNA复制的各种蛋白质和酶的鉴定,主要是利用DNA复制突变体的互补和体外复制系统的重建两种方法。互补的原理是利用某种野生型的蛋白质,去恢复特定的DNA复制缺陷突变体的复制功能,从而确定参与复制的蛋白质。重建的原理是在较为简单的体外复制系统(如SV40病毒复制系统)中,先人为去掉某种成分,致使复制不能正常进行,然后在复制系统中,逐一添加分步收集的可能参与复制的蛋白质抽取物,检测能否恢复复制活性,从而确定复制蛋白。有时,添加的蛋白质可能来自于其他物种,这样可以从其他物种中,找到同源或同工的蛋白质。为了方便理解重建的原理,这里可以打一个比方加以说明。假定你的一台电脑坏了一个部件而不能运转,那么如何迅速找到是哪一个部件出了故障呢?这时可以用类似重建的手段来确定:首先将一台运转正常的电脑的各个部件拆开,显然来自这台正常电脑内所有的部件都应该是正常的(相当于野生型蛋白质)。然后,将坏掉的电脑逐一取出一个部件(如内存条或主板),再用正常电脑的相应部件取而代之。如果某一个部件经过替换后,坏的电脑恢复正常了,这就等于找到了坏的部件(相当于突变型蛋白质)。这两种方法对于参与其他过程(如信号转导、转录、转录后加工、翻译、细胞周期的调控等)的蛋白质的鉴定也很有帮助。例如,为了找到人细胞内参与细胞周期的某种蛋白质,可先将酵母细胞内一种与细胞周期有关的蛋白质突变,这样的酵母细胞周期肯定会有异常。然后,将人细胞内各种可能与细胞周期有关的蛋白质,导入到突变的酵母细胞中。如果其中的某一组分加入以后,酵母的细胞周期恢复正常,那么被加入的蛋白质极有可能就是人细胞内与细胞周期有关的蛋白质。

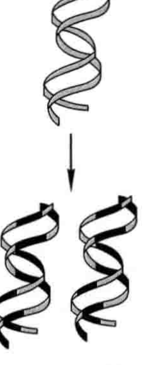
第三十三章 DNA 复制

早在 1945 年,生物学家已有足够的证据证明 DNA 是生物最主要的遗传物质。作为遗传物质的 DNA 至少应该具有两个最基本的功能:首先是高度精确的复制能力——这是一种生物能将其遗传信息准确、稳定地进行传递的前提。其次是具有编码蛋白质和其他生物分子的能力——这是细胞行使其全部功能的基础。

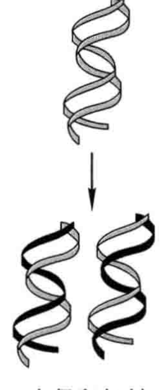
本章只介绍 DNA 的第一个功能,将重点讨论 DNA 复制的一般特征、参与 DNA 复制的主要酶和蛋白质的结构与功能以及几种具有代表性的复制模式,包括细菌染色体 DNA 的"θ 复制"、真核细胞核的 DNA 复制、线粒体 DNA 的"D 环复制"、某些病毒的"滚环复制"和古菌的 DNA 复制,并对 DNA 复制过程中表现出来的高度忠实性进行总结。DNA 的第二个功能将在 DNA 转录和翻译这两章中作详细的介绍。



全保留复制



弥散性复制



半保留复制

图 33-1 DNA 复制可能的三种方式

第一节 DNA 复制的一般特征

DNA 复制既可以发生在原核细胞的细胞质,也可以发生在真核细胞的细胞核、叶绿体或线粒体。不同的复制系统具有一些共同的特征,它们包括:

(1) 以亲代 DNA 的两条母链作为模板(template),四种 dNTP 为前体,还需要 Mg^{2+} 根据 DNA 的双螺旋结构模型,构成双螺旋的两条互补链可各自作为另外一条链的模板,按照碱基互补配对规则,复制产生新的 DNA 分子。在复制中,dATP、dGTP、dCTP 和 dTTP 作为构建新的多聚脱氧核苷酸的原料。

(2) 作为模板的 DNA 需要解链 作为模板的碱基序列隐藏在双螺旋结构的核心,因此,DNA 在复制的时候,必须进行解链,才能暴露出内部的模板序列,同时,游离出原来碱基对之间的氢键供体或受体,为建立新的互补碱基对创造条件。

(3) 半保留(semi-conservative)复制 Watson 和 Crick 在 1953 年提出著名的 DNA 双螺旋结构模型的同时,对 DNA 复制可能的模式曾做了大胆的预测。他们认为,DNA 复制可能采取的是一种半保留方式,即复制的时候,亲代 DNA 的两条母链先解链分离,然后分别作为新链合成的模板。在最终得到的两个子代 DNA 分子中,一条链是新合成的子链,另一条链是原来的母链,换句话说,原来作为模板的两条 DNA 母链被半保留在子代的 DNA 分子之中。然而,理论上讲,DNA 复制还可能采取另外两种方式,即全保留方式和弥散性(dispersive)方式。在全保留方式中,亲代的两条 DNA 链被完全保留在一个子代的 DNA 分子之中,而在弥散性复制中,来自亲代的 DNA 链弥散性地分布在子代 DNA 的任意一条链之中(图 33-1)。

1958 年,Matthew Meselson 和 Franklin Stahl 用实验证明,细胞内 DNA 的复制以半保留方式进行。该实验的设计思路极为巧妙,被誉为生物学最美丽的实验(详见本章科学故事),整个实验流程参见图 33-2,具体步骤是:先将大肠杆菌置于 $^{15}NH_4Cl$ 为唯一 N 源的培养基上连续培养十多代,以使胞内 DNA 分子上所有的 N 原子都成为较重的 ^{15}N;然后,从上述培养基中收集细菌,其中的一部分用于抽取 DNA,另一部分改放在以 $^{14}NH_4Cl$ 为 N 源的培养基中继续培养。在将不同培养代数的大肠杆菌进行收集、裂解和 DNA 抽取后,用 CsCl 密度梯度离心的方法,分析各代 DNA 在离心管中的区带位置,并与一直在 ^{15}N 或 ^{14}N 培养基中培养的大肠杆菌(对照)的 DNA 区带位置进行比较,结果发现"0 代"DNA 为 1 条高密度带(DNA 2 条链上的 N 原子全部是 ^{15}N,因此两条链都是重链,即 H/H-DNA,H 为 heavy

对照:一直在 ^{14}N 培养基中生长的大肠杆菌

连续多代在 ^{15}N 培养基上生长的大肠杆菌(F_0代)

^{15}N

转移到 ^{14}N 培养基上继续培养

F_1代

F_2代

等

裂解细胞,抽取DNA

CsCl密度梯度离心

$^{14}N^{14}N$

$^{15}N^{15}N$

$^{15}N^{14}N$

$^{14}N^{14}N$

$^{15}N^{14}N$

图 33-2 Meselson 和 Stahl 的证明 DNA 半保留复制的实验流程(Hartwell 等,2011)

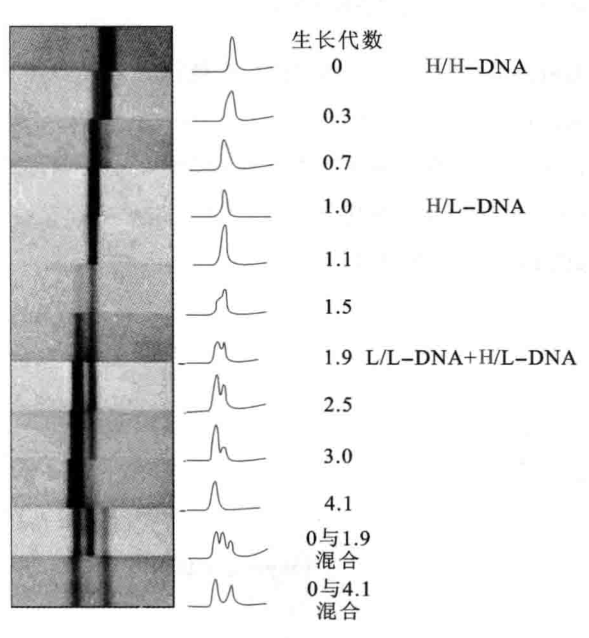


图 33-3 Meselson 和 Stahl 的实验结果

的缩写),“第 1 代”DNA 得到 1 条中密度带(一条链为重链,另一条链为轻链,即 H/L-DNA,L 为 light 的缩写),而“第 2 代”DNA 有中密度(H/L-DNA)和低密度(L/L-DNA)两条带。这样的结果(图 33-3)与大肠杆菌 DNA 半保留复制的预期结果完全一致,因此,有理由相信,至少大肠杆菌的 DNA 复制是半保留复制。

(4) 需要引物 DNA 复制不能从头合成,只能在事先合成好的引物上进行链的延伸。作为引物的主要是短的 RNA,其长度一般为 6~15 nt,少数是蛋白质。但如果在体外进行 DNA 复制,如 PCR 和双脱氧法测序,就需要使用人工合成的 DNA 作为引物。

(5) 复制的方向始终是 5′→ 3′。

(6) 具有固定的起点 体内的 DNA 复制并不是随机启动的,而是从特定的区域开始,即具有固定的起点。作为复制起点的碱基序列被称为复制起始区(origin)。

复制起始区通常具有三个共同的特征:①由多个短的重复序列组成;②能够被多亚基的复制起始区结合蛋白识别;③通常富含 AT 碱基对,这显然有利于 DNA 复制启动时的解链,因为 AT 碱基对含有的氢键数目低于 GC 碱基对。

Quiz1 如果将 Meselson 的实验步骤改成:先在 ^{14}N 的培养基中培养,然后后面连续几代改放在 ^{15}N 的培养基上培养,并用密度梯度离心分析各代 DNA 的条带变化,你认为这样仍然能成功证明 DNA 半保留复制吗?

Quiz2 你认为为什么 DNA 复制需要 RNA 引物?

Quiz3 试设计一个实验,证明 DNA 复制方向总是从 5′→ 3′。



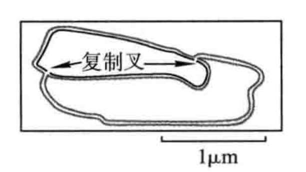


图 33-4 复制叉的结构

细菌 DNA 复制起始区通常只有一个，而真核生物有多个复制起始区，许多古菌也具有多个复制起始区。每一个复制起始区构成一个最小的复制单位——复制子(replicon)，因此细菌整个染色体 DNA 就是一个复制子，而真核生物和许多古菌的每一个染色体 DNA 就有几个或多个复制子。

当 DNA 复制从起始区启动的时候，起始区因发生解链而形成叉状结构(图 33-4)，这样的结构称为复制叉(replication fork)。

(7) 多为双向复制，少数为单向复制　绝大多数 DNA 的复制在起始区启动以后，便“左右开弓”，向两个方向同时展开，进行双向复制(bidirectional)。少数 DNA 的复制只能向一个方向进行，即单向复制。进行双向复制的 DNA 在复制起始区形成 2 个复制叉，而单向复制的 DNA 只有 1 个复制叉。

(8) 半不连续性(semi-discontinuous)　由于构成 DNA 双螺旋的两条链呈反平行，而 DNA 复制的方向只能是 5′→3′，因此，一个复制叉内发生的 DNA 复制极有可能采取的是一种半不连续的模式，即其中的一条子链连续合成，与复制叉前进的方向相同，而另一条子链则不连续合成，与复制叉前进的方向相反。不连续合成是指先合成一些小的不连续的片段，然后再将这些不连续的片段连接起来。

首先提出半不连续复制的是 Reiji Okazaki。为了确定大肠杆菌的染色体 DNA 复制是不是以这种方式进行，1958 年，Okazaki 设计了脉冲标记(pulse labeling)和脉冲追踪(pulse chase)的实验，其中脉冲标记的目的在于即时标记在特定时段内合成的 DNA，而脉冲追踪的目的则是要弄清那些被标记上的 DNA 片段后来的去向。他以大肠杆菌染色体的 DNA 复制为研究对象，实验流程参看图 33-5。结果表明，染色体 DNA 的复制至少有一条子链是先合成短的 DNA 片段，再形成较长的分子(图 33-6)。

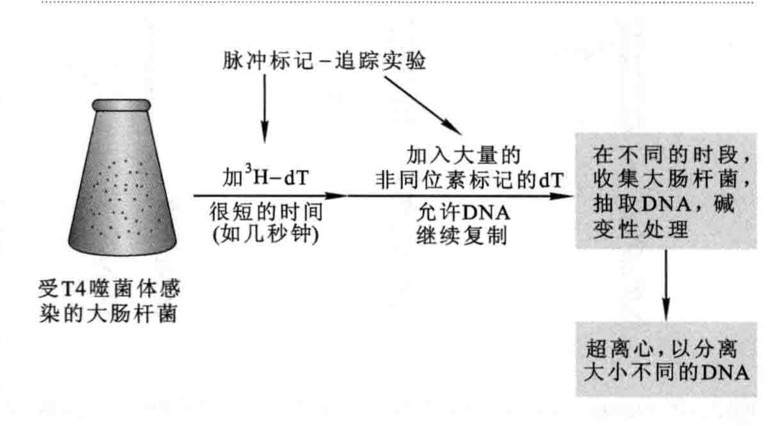


图 33-5 Okazaki 的脉冲标记和脉冲追踪的实验

为了纪念 Okazaki 在 DNA 复制研究中的杰出贡献，人们将在复制叉中不连续合成的 DNA 片段称为冈崎片段(Okazaki fragments)。而按照 Okazaki 最初的建议，连续合成的 DNA 子链称为前导链(leading strand)，不连续合成的子链称为后随链(lagging strand)(图 33-7)。

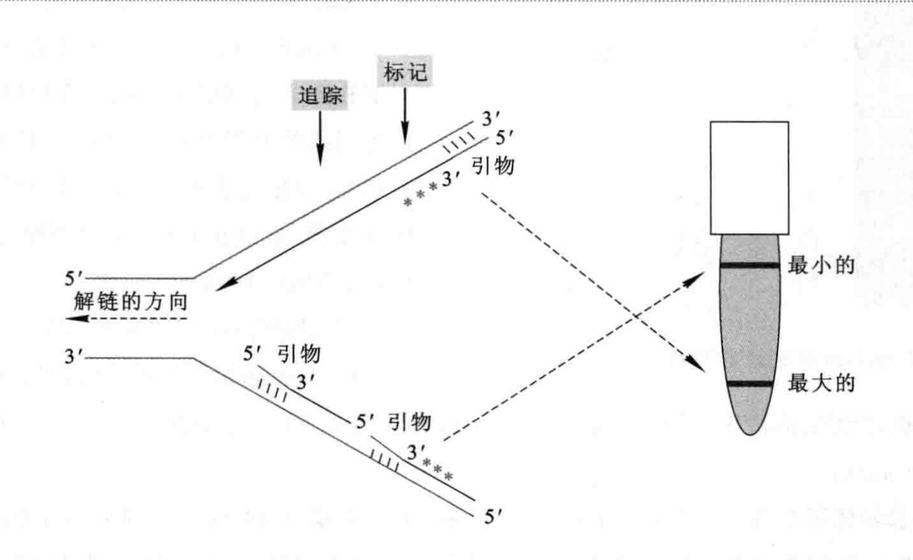


图 33-6 Okazaki 的脉冲标记和脉冲追踪的实验结果分析

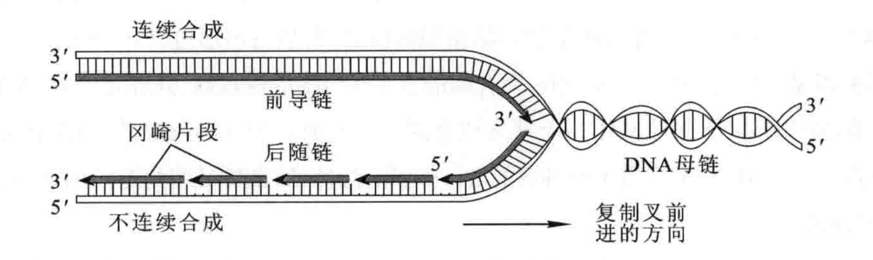


图 33-7　DNA 的半不连续复制

(9) 具有高度的忠实性　DNA 复制出错的概率很小,其忠实性明显高于转录、逆转录、RNA 复制和翻译。之所以 DNA 复制具有高度的忠实性,是因为细胞内存在着一系列互补的校对和纠错机制。

(10) 具有高度的进行性　DNA 复制的进行性(processivity)是指催化 DNA 复制的聚合酶从与模板结合到与模板解离的这段时间内,所催化参入到 DNA 子链上的核苷酸数目。细胞内有多种进行性不同的 DNA 聚合酶,但参与 DNA 复制的一定是进行性高的。

Quiz4 如何设计一个实验测定一种 DNA 聚合酶的进行性?

第二节　参与 DNA 复制的主要酶和蛋白质

DNA 复制是在一系列酶和蛋白质的催化或协助下完成的,其中涉及的主要酶和蛋白质有 DNA 聚合酶、DNA 解链酶、单链结合蛋白、DNA 拓扑异构酶、DNA 引发酶、切除引物的酶、DNA 连接酶、尿嘧啶-DNA 糖苷酶和端聚酶等,现分别介绍如下。

一、DNA 聚合酶

DNA 聚合酶的全名是依赖于 DNA 的 DNA 聚合酶(DNA-dependent DNA polymerase),顾名思义就是以 DNA 为模板,催化 DNA 合成的聚合酶。

DNA 聚合酶是参与 DNA 复制的主要酶,其催化的反应通式为:

$$\text{引物-OH} + (\text{dNTP})_n \xrightarrow[\text{DNA 模板、}Mg^{2+}]{\text{DNA 聚合酶}} \text{引物-O-dNMP} + (\text{dNTP})_{n-1} + \text{PPi}$$

反应中形成的焦磷酸在焦磷酸酶的催化下迅速水解,使得聚合反应趋于完全。该酶的许多性质直接决定了 DNA 复制的一些基本特征。例如,该酶不能催化 DNA 的从头合成决定了 DNA 复制需要引物,而它只能从 5′→3′催化聚合反应决定了 DNA 复制的单向性。

(一) 细菌的 DNA 聚合酶

已在大肠杆菌细胞中发现 5 种 DNA 聚合酶,即 DNA 聚合酶Ⅰ、Ⅱ、Ⅲ、Ⅳ和Ⅴ,现分别加以讨论。

1. 大肠杆菌的 DNA 聚合酶Ⅰ

DNA 聚合酶最早由 Arthur Kornberg 在大肠杆菌中发现。在 1953 年 Watson 和 Crick 提出 DNA 双螺旋模型并对 DNA 复制的可能机制做出预测以后,许多生物学家在苦苦寻找那种能直接催化 DNA 复制的酶。1957 年,Kornberg 在大肠杆菌中分离得到第一种能够在体外催化 DNA 合成的酶,该酶不久被命名为 DNA 聚合酶Ⅰ,以区别后来发现的几种 DNA 聚合酶。DNA 聚合酶Ⅰ有时也被称为 Kornberg 酶。

在对 DNA 聚合酶Ⅰ进行更为详尽的研究以后发现,此酶所具有的一些性质大大出乎人们的意料。该酶只由 1 条肽链组成,除了具有 5′→3′的 DNA 聚合酶活性以外,还具有 5′-外切酶和 3′-外切酶的活性。

DNA 聚合酶Ⅰ为什么还同时具有核酸外切酶活性呢?原来 3′-外切酶活性是被聚合酶用来自我校对的。聚合酶活性并不能保证在复制中参入到 DNA 3′端的碱基 100% 正确。关键在于倘若出现

Quiz5 你如何证明 3′-外切酶活性参与 DNA 复制的校对?

错配的碱基，细胞有没有补救的措施。实际上，当正在复制的 DNA 子链的 3′端出现错配碱基的时候，DNA 聚合酶凭借其内在的 3′－外切酶活性切除错误的核苷酸，然后再通过其 5′→ 3′的聚合酶活性换上正确配对的碱基。实验证明，一种 DNA 聚合酶的 3′－外切酶活性越高，其催化的 DNA 复制的忠实性就越高。如果一种 DNA 聚合酶本来就缺乏或者因为突变而丧失这种酶活性，就会导致 DNA 复制忠实性的下降。至于 DNA 聚合酶 I 所具有的 5′－外切酶活性，后来被证明是专门用来切除存在于 5′端的 RNA 引物的。

Hans Klenow 使用枯草杆菌蛋白酶或胰蛋白酶处理 DNA 聚合酶 I，结果得到大小两个片段：大片段被称为 Klenow 片段或 Klenow 酶，它含有大小两个结构域，其中小结构域具有 3′－外切酶活性，大结构域具有 5′→ 3′聚合酶活性；小片段只有 5′－外切酶活性。

1987 年，Tom Steitz 等人得到了 Klenow 酶的晶体结构。X 射线衍射数据表明，该酶有“校对活性”和“聚合酶活性”两种构象。图 33–8 为处于“校对活性状态下”的 Klenow 酶的三维晶体结构，其外形像右手，分子表面含有两个近乎垂直的裂缝。其中有一个在大结构域的表面，位于由 L–P 螺旋组成的手指(finger)与由 H 螺旋和 I 螺旋组成的拇指（thumb）之间，含有与单链 DNA 模板结合的位点，而聚合酶活性中心位于裂缝的底部。该裂缝宽约 2.2 nm，长约 3 nm，大小和形状正适合结合 B–DNA。外切酶活性中心位于拇指和手指之间的掌心上，与聚合酶相距仅 3.5 nm。另外一个裂缝为双链 DNA 结合位点，与双链 DNA 的结合仅涉及与磷酸骨架接触，并不涉及碱基，这与聚合酶与 DNA 结合无序列特异性是一致的。

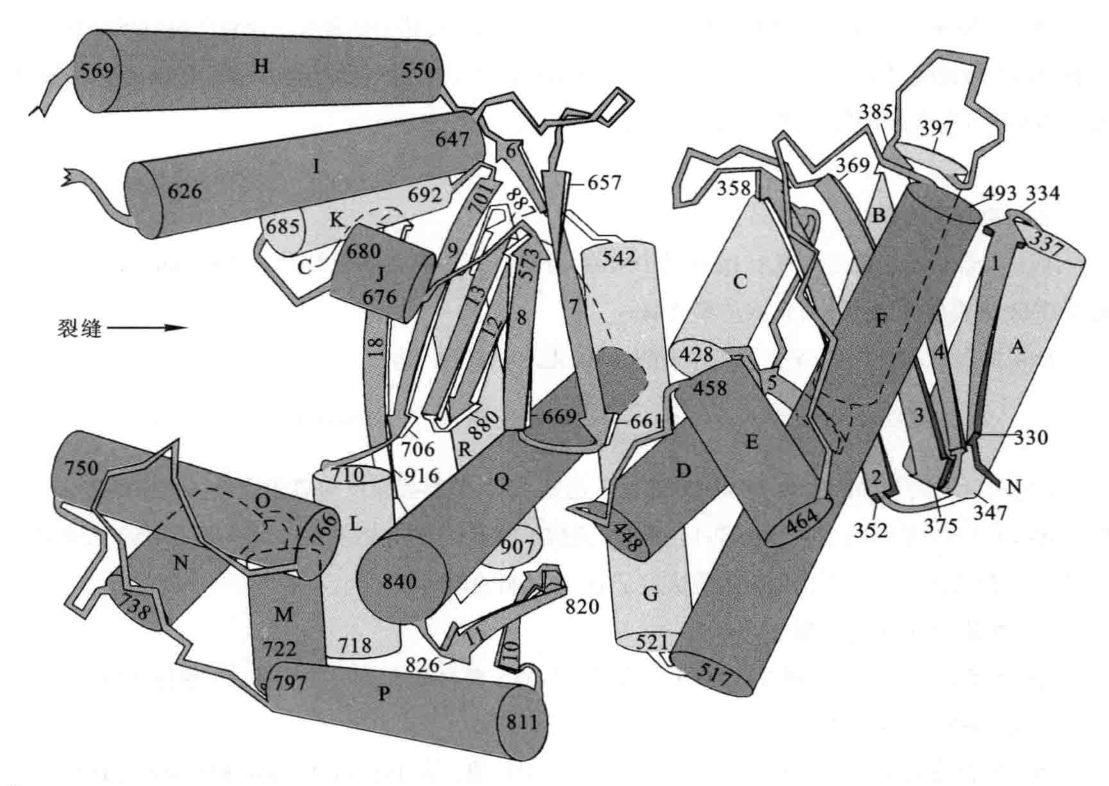


图 33–8　Klenow 酶的晶体结构模型（Weaver，2011）

如图 33–9 所示，当 DNA 复制在子链的 3′端出现错配的时候，错配的碱基很容易发生解链而落入 3′－外切酶的活性中心被切除。丢掉错配碱基的 3′端再回到聚合酶活性中心，使 DNA 链恢复延伸。当酶处于“聚合酶活性状态”时，位于聚合酶活性中心的一些高度保守的碱性氨基酸残基（如 Arg 754、Arg 682、Lys 758 和 His 734）直接与进入活性中心的 dNTP 上的磷酸基团相互作用。除了这些带正电荷的保守性氨基酸残基以外，进入活性中心的 dNTP 还与一些带负电荷的氨基酸残基（如 Asp 882、Asp 705 和 Glu 883）发生作用。

DNA 复制首先需要合成 RNA 引物。引物合成好以后，并与模板链共同诱导 DNA 聚合酶在外切

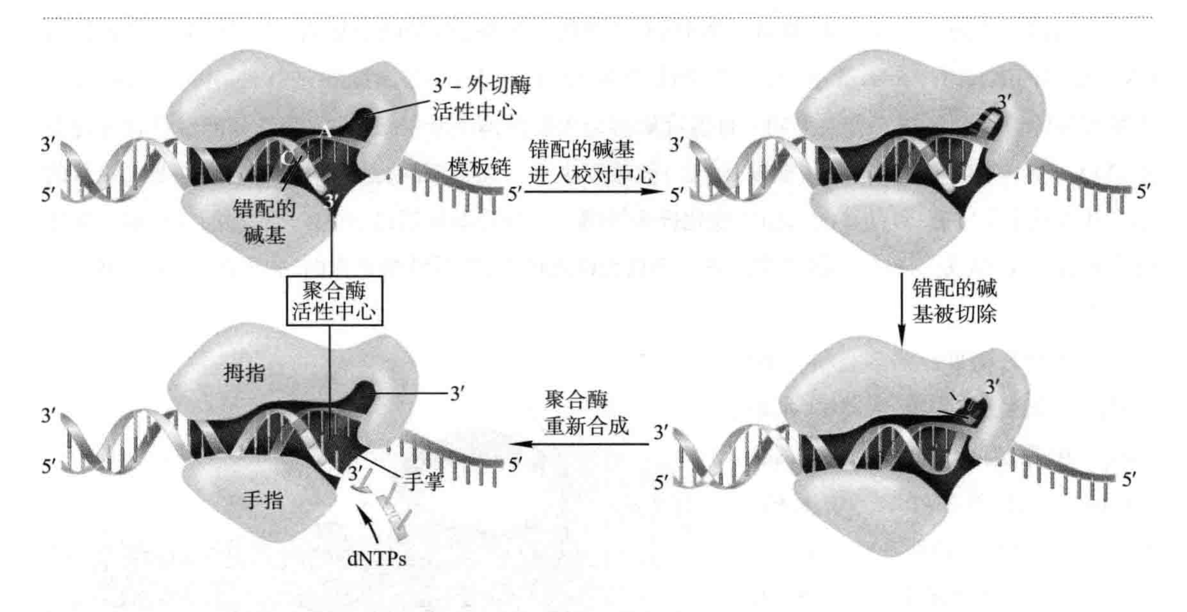


图 33-9　DNA 聚合酶的聚合和校对(Brooker,2009)

酶活性中心附近形成一个新的裂缝。随后,RNA 引物和模板链从这个新的裂缝进入聚合酶活性中心,DNA 的复制由此在引物的 3′端被启动。

在 DNA 聚合酶Ⅰ被发现以后的很长一段时间内,人们曾误以为它是大肠杆菌唯一的 DNA 聚合酶。然而,此酶的一些性质(如速率太慢、酶量太多和进行性不够高)显示,它并不适合催化大肠杆菌染色体 DNA 的复制,其中最重要的证据来自遗传学突变实验:1969 年 De Luca 和 Cairns 报道,一种缺乏聚合酶Ⅰ活性的大肠杆菌突变株照样能够生存,可进行正常的 DNA 复制,但对各种诱变剂(如紫外线)的作用更为敏感。这种突变株的发现彻底否定了 DNA 聚合酶Ⅰ为催化大肠杆菌染色体 DNA 复制的主要聚合酶,而该突变株自然就成为科学家寻找其他 DNA 聚合酶的"宝地"。不出所料,另外两种 DNA 聚合酶即聚合酶Ⅱ和Ⅲ很快在这种突变株细胞中相继被发现。

Quiz6　如果将 DNA 聚合酶Ⅰ敲除,你认为大肠杆菌还能生存吗?为什么?

2. 大肠杆菌 DNA 聚合酶Ⅱ

此酶也同时具有 5′→ 3′聚合酶活性和 3′－外切酶活性,但无 5′－外切酶活性。缺乏此酶活性的突变株在细胞生长和 DNA 复制上没有任何缺陷。由于它催化的聚合反应速率特别慢,难以满足体内染色体 DNA 复制的需要,此酶最有可能参与 DNA 的修复。

3. 大肠杆菌 DNA 聚合酶Ⅲ

DNA 聚合酶Ⅲ由多个亚基组成(表 33-1),虽然也具有 5′→ 3′聚合酶活性和 3′－外切酶活性,但却分属不同的亚基。

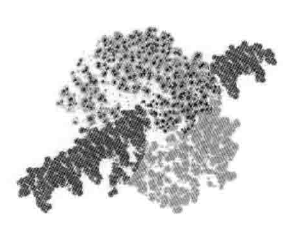
表 33-1　大肠杆菌 DNA 聚合酶Ⅲ

			亚基	功能
全酶	polⅢ′	核心酶	α	5′→ 3′聚合酶活性
			ε	3′→ 5′外切酶活性
			θ	α 和 ε 的装配
			τ	将全酶装配到 DNA
			β	滑动钳(进行性因子)
			γ	滑动钳装载复合物
			δ	滑动钳装载复合物
			δ′	滑动钳装载复合物
			χ	滑动钳装载复合物
			ψ	滑动钳装载复合物

该酶被认为是参与大肠杆菌染色体 DNA 复制的主要酶，相关的证据有：①酶的 V_{max} 接近体内 DNA 复制的实际值；②酶量适中，一个细胞约有 10 个酶分子；③高度的进行性，与实际值相近；④最直接的遗传学证据。有人得到一种温度敏感型大肠杆菌突变株，这种突变株只能生存在允许温度（30℃）下，当温度上升到限制温度（45℃）就难以生存。究其原因，是因为编码聚合酶Ⅲα 亚基的 *polC* 基因发生了突变，致使该酶对温度变化极为敏感。当环境温度超过 30℃以后，此酶就很容易变性而丧失活性，DNA 复制也就不能正常进行。而在允许温度下，酶活性是正常的，胞内的 DNA 复制也就很正常。



切面观



正面观

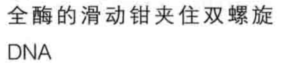
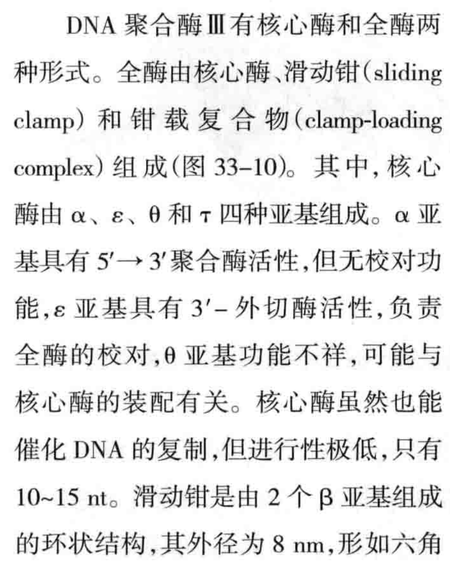


图 33-11　DNA 聚合酶Ⅲ全酶的滑动钳夹住双螺旋 DNA



DNA 聚合酶Ⅲ有核心酶和全酶两种形式。全酶由核心酶、滑动钳（sliding clamp）和钳载复合物（clamp-loading complex）组成（图 33-10）。其中，核心酶由 α、ε、θ 和 τ 四种亚基组成。α 亚基具有 5′→3′聚合酶活性，但无校对功能，ε 亚基具有 3′-外切酶活性，负责全酶的校对，θ 亚基功能不详，可能与核心酶的装配有关。核心酶虽然也能催化 DNA 的复制，但进行性极低，只有 10~15 nt。滑动钳是由 2 个 β 亚基组成的环状结构，其外径为 8 nm，形如六角形，内部为一空洞，直径为 3.5 nm，大于 B-DNA 或 A-DNA 的双螺旋直径。在 DNA 复制的时候，这种钳状结构能松散地夹住 DNA 模板（图 33-11），并自由地向前滑动，这就大大提高了聚合酶Ⅲ的进行性。钳载复合物由其他几种亚基组成，其中的 γ 亚基具有 ATP 酶活性，其功能是以 ATP 水解为动力，打开钳子，帮助滑动钳装载到 DNA 模板上。

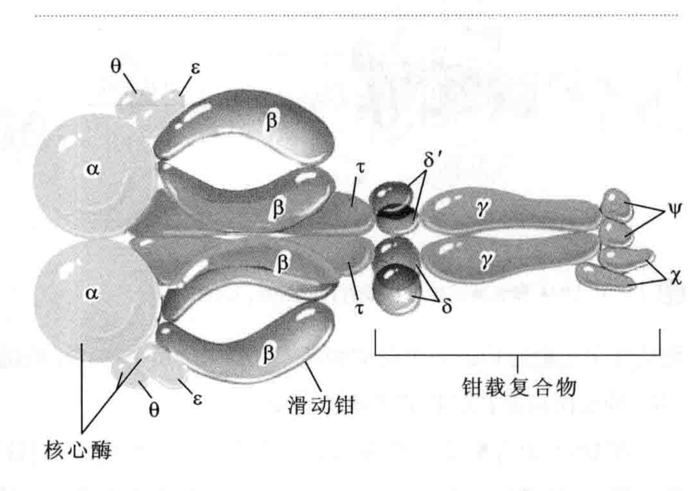


图 33-10　大肠杆菌 DNA 聚合酶Ⅲ全酶的结构模型（Horton，2006）

► 表 33-2　DNA 聚合酶Ⅰ、Ⅱ和Ⅲ的性质和功能

性质	DNA 聚合酶Ⅰ	DNA 聚合酶Ⅱ	DNA 聚合酶Ⅲ
结构基因	*polA*	*polB*	*polC*
相对分子质量 / $\times 10^3$	103	90	130
分子数 / 细胞	400	100	10
V_{max}（参入的 nt/s）	16~20	2~5	250~1 000
3′- 外切酶活性	√	√	√
5′- 外切酶活性	√	×	×
进行性 /nt	3~200	10 000	500 000
突变体表型	UV 敏感、硫酸二甲酯敏感	无	DNA 复制温度敏感型
生物功能	DNA 修复、RNA 引物切除	DNA 修复	染色体 DNA 复制

4. 大肠杆菌 DNA 聚合酶Ⅳ和Ⅴ

大肠杆菌 DNA 聚合酶Ⅳ和Ⅴ直到 1999 年才被发现，都属于易错的 DNA 聚合酶，参与 DNA 的修复合成，特别是 SOS 修复过程中的跨损伤合成（参看第三十四章“DNA 的损伤、修复和突变”）。

（二）真核生物的 DNA 聚合酶

到目前为止，已在真核生物的细胞中发现至少 15 种不同的 DNA 聚合酶，除了 5 种发现较早的 DNA 聚合酶 α、β、γ、δ 和 ε 以外（表 33-3），还有聚合酶 θ、ζ、η、κ、ι、μ、λ、ψ 和 ξ 等，这些新发现的 DNA 聚合酶除了 θ 以外，一般无 3′- 外切酶的活性，因此无校对功能。它们主要参与 DNA 的跨越合成，以克服损伤对 DNA 复制的不利影响。

表 33-3 真核细胞 DNA 聚合酶 α、β、γ、δ 和 ε 的性质和功能

性质	DNA 聚合酶 α	DNA 聚合酶 β	DNA 聚合酶 γ	DNA 聚合酶 δ	DNA 聚合酶 ε
亚细胞定位	细胞核	细胞核	线粒体基质	细胞核	细胞核
引发酶活性	有	无	无	无	无
亚基数目	4	1	4	2	≥4
内在的进行性	中等	低	高	低	高
在 PCNA 存在时的进行性	中等	低	高	高	高
3′-外切酶活性	×	×	√	√	√
5′-外切酶活性	×	×	×	×	×
生物功能	细胞核 DNA 复制	细胞核 DNA 修复	线粒体 DNA 复制	细胞核 DNA 复制	细胞核 DNA 复制和修复

表 33-3 中的 PCNA 表示分裂细胞核抗原(proliferating cell nuclear antigen)为聚合酶 δ 的辅助蛋白，由三个亚基组成，其功能类似于大肠杆菌 DNA 聚合酶Ⅲ的 β 亚基，在真核 DNA 复制中作为滑动钳以提高该酶的进行性(图 33-12)。

二、DNA 解链酶

DNA 解链酶(DNA nelicase，也称 DNA 解螺旋酶)是一类催化 DNA 双螺旋解链的酶。细胞内所有涉及解链的过程都需要解链酶，如 DNA 复制、转录、修复和重组。

任何一种 DNA 解链酶都能结合 DNA，这种结合与 DNA 序列无关。大多数解链酶优先结合 DNA 的单链区域，少数解链酶优先结合 DNA 的双链区域。

解链酶除了能结合 DNA 以外，还能结合 NTP(主要是 ATP)，并同时具有依赖于 DNA 的 NTP 酶活性。解链酶需要 NTP 的水解为解链提供能量。

所有的 DNA 解链酶都具有移位酶活性，只有这样，才能沿着被结合的 DNA 模板向一侧移动，并不断地解开 DNA 双链区域。



图 33-12 真核细胞 DNA 聚合酶 δ 由 PCNA 三个亚基构成的滑动钳结构

三、单链 DNA 结合蛋白

图 33-13 大肠杆菌 DnaB 蛋白的解链酶活性

单链 DNA 结合蛋白(single-stranded DNA-binding proteins，SSB)是一种专门与 DNA 单链区域结合的蛋白质，为 DNA 复制、修复和重组所必需。SSB 本身并没有任何酶的活性，但通过与 DNA 单链区段的结合，在 DNA 复制、修复和重组中至少有 3 个作用：①暂时维持 DNA 的单链状态，防止互补的单链在作为复制模板之前重新退火成双螺旋；②防止 DNA 的单链区域自发形成链内二级结构(如链内双螺旋)而影响 DNA 聚合酶的进行性；③防止核酸酶对 DNA 单链区域的水解。

四、DNA 拓扑异构酶

拓扑异构酶(topoisomerase)是一类通过催化 DNA 链的断裂、旋转和重新连接而直接改变 DNA 拓扑学性质的酶。此类酶不但可以解决在 DNA 复制、转录、重组和染色质重塑过程中遇到的拓扑学障碍，

而且能够细调细胞内 DNA 的超螺旋程度,以促进 DNA 与蛋白质的相互作用。同时还可以防止 DNA 形成有害的过度超螺旋。

到目前为止,已发现所有拓扑异构酶的作用都是通过两次转酯反应来完成的(图 33-14):第一次转酯反应由活性中心 1 个 Tyr 残基上的氧原子亲核进攻 DNA 链上的 3′,5′-磷酸二酯键,形成以磷酸酪氨酸酯键相连的酶与 DNA 的共价中间复合物。与此同时,DNA 链发生断裂。形成的这种共价中间复合物既固定了被断裂的磷酸二酯键中的能量,又大大降低了 DNA 链上出现异常的永久性切口的可能性。在断裂的 DNA 链进行重新连接之前,DNA 的另一条链或另一个 DNA 双螺旋通过切口,从而导致其拓扑学结构发生变化。最后,发生第二次转酯反应,这一次转酯反应可视为第一次转酯反应的逆反应,由 DNA 链断裂处的自由羟基亲核进攻酶与 DNA 之间的磷酸酪氨酸酯键,致使原来断裂的 3′,5′-磷酸二酯键重新形成,而酶则恢复到原来的状态。

e33-1 DNA 拓扑异构酶的亚类及性质比较

按照 DNA 链的断裂方式,拓扑异构酶被分为两种类型,即Ⅰ型和Ⅱ型,而每种形式还有 A 类和 B 类。不同类型的拓扑异构酶在催化机制、对 Mg^{2+} 和 ATP 的依赖性以及对超螺旋连环数的改变不尽相同。Ⅰ型拓扑异构酶在作用过程中,只能切开 DNA 的一条链,而Ⅱ型会同时交错切开 DNA 的两条链(图 33-15),并在消耗 ATP 的同时,将一个 DNA 双螺旋从一处经过另一个双螺旋的裂口“主动转运”到另外一处。

Quiz7 Ⅱ型拓扑异构酶每作用一次,DNA 超螺旋的连环数有何变化?

Ⅱ型拓扑异构酶主要参与 DNA 复制,既可以在 DNA 分子中引入有利于复制的负超螺旋,又可以及时清除复制叉前进中形成的正超螺旋,还能分开复制结束后缠绕在一起的两个子代 DNA 分子。其催化的反应依赖于 ATP。在 ATP 的存在下,一个 DNA 双螺旋上的两条链同时出现切口。随后,另一个 DNA 双螺旋穿过切口。最后,切口重新连接。在其间,可完成几种不同类型的拓扑学转变,包括松弛正、负超螺旋,环形 DNA 的连环化(catenation)和去连环化(decatenation)。

细菌的旋转酶(gyrase)就属于Ⅱ型拓扑异构酶,由 2 个 A 亚基和 2 个 B 亚基组成,参与细菌染色体 DNA 的复制。在消耗 ATP 的条件下,它可在共价闭环 DNA 分子中连续引入负超螺旋。在无 ATP 的情况下,可以松弛负超螺旋。环丙沙星(ciprofloxin)和新生霉素(novobiocin)是两种作用旋转酶的抗

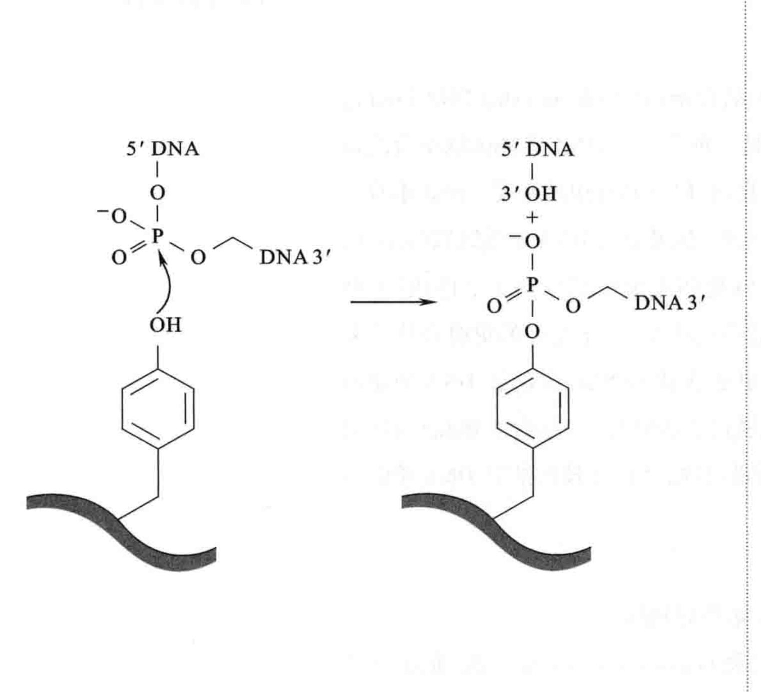


图 33-14 DNA 拓扑异构酶的催化的转酯反应

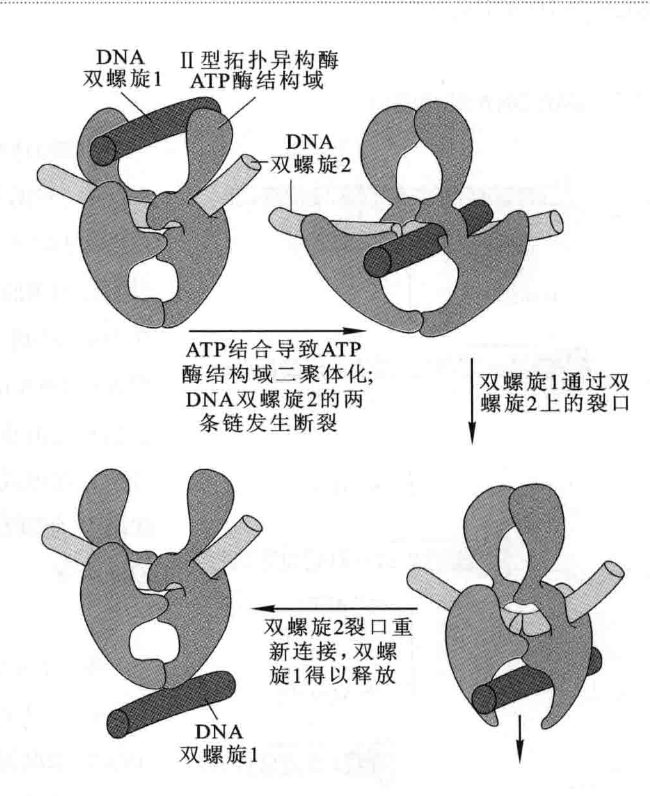


图 33-15 Ⅱ型 DNA 拓扑异构酶的作用机制(Alberts 等,2007)

生素,前者的作用位点是 A 亚基,能抑制旋转酶的 ATP 酶活性,后者的作用位点是 B 亚基,能增强旋转酶切断 DNA 链的能力,但抑制 DNA 链的重新连接。

五、DNA 引发酶

DNA 引发酶(DNA primase)催化 RNA 引物的合成,在启动引物合成的时候与复制叉结合,在引物合成好以后又立刻与复制叉解离。由于 DNA 复制的半不连续性,引发酶在前导链上只需要引发一次,而在后随链上则需要引发多次。

大肠杆菌的引发酶由 *dnaG* 基因编码,在体内催化合成 11 ± 1 nt 长的 RNA 引物。真核细胞的引发酶在体内与 DNA 聚合酶 α 紧密地结合在一起,催化合成 7~10 nt 长的 RNA 引物。

六、切除引物的酶

RNA 引物只是用来启动 DNA 的复制,迟早要被切除。实际上,细胞内有专门的酶负责切除 RNA 引物。细菌切除 RNA 的酶是 DNA 聚合酶 I 或核糖核酸酶 H(RNase H),其中 DNA 聚合酶 I 凭借自带的 5′-外切酶活性,切除总是位于 5′端的引物,RNase H 专门水解与 DNA 杂交的 RNA,包括 RNA 引物。真核细胞没有哪一种 DNA 聚合酶兼有 5′-外切酶活性,因此切除 RNA 引物的酶不可能是 DNA 聚合酶,而是利用 RNase HI/FEN1。FEN1 被称为翼式内切酶(flap endonuclease),具有 5′-外切酶和内切酶活性。一般认为,RNase HI 负责切割连接在 DNA 链 5′端的 RNA,但会留下一个核糖核苷酸,由 FEN1 最后来切除。

七、DNA 连接酶

DNA 连接酶(DNA ligase)能够催化一个双螺旋 DNA 分子内相邻核苷酸的 3′-羟基和 5′-磷酸,甚至两个双螺旋 DNA 分子两端的 3′-羟基和 5′-磷酸,发生连接反应,形成 3′,5′-磷酸二酯键。这一类酶主要参与 DNA 复制,也参与 DNA 修复和重组。

DNA 连接酶在 DNA 复制中的作用是“缝合”相邻的冈崎片段,使不连续合成的后随链成为一条连续的链,而在 DNA 修复和重组中的作用则是闭合 DNA 链上产生的切口。

DNA 连接酶在催化连接反应时需消耗能量。根据能量供体的性质,连接酶可以分为两类:第一类使用 NAD^+,第二类使用 ATP。绝大多数细菌的 DNA 连接酶属于第一类,真核生物、古菌、病毒和少数细菌的连接酶属于第二类。

Quiz8 在大量 AMP 存在的情况下,DNA 连接酶可切开 DNA,为什么?

所有的 DNA 连接酶催化的反应都可分为 3 步(图 33-16):①连接酶活性中心的一个 Lys 残基的 ε-氨基亲核进攻 ATP 或 NAD^+ 的磷酸酯键,形成共价的酶 -AMP 中间物,同时释放出 PPi 或 NMN;② AMP 被转移到 DNA 链切口处的 5′-磷酸上;③切口处的 3′-OH 进攻 AMP-DNA 之间的键,致使切口处相邻的核苷酸之间形成 3′,5′-磷酸二酯键,同时释放出 AMP。

Quiz9 DNA 连接酶和拓扑异构酶都使用共价催化机制吗?

八、尿嘧啶 -DNA 糖苷酶

尿嘧啶 -DNA 糖苷酶专门用来切除出现在 DNA 链上的 U。一般说来,U 出现在 DNA 链上有两个原因:一是在 DNA 复制过程中,细胞中存在的 dUTP 代替 dTTP 直接进入新合成的 DNA 链;二是 DNA 分子中的 C 发生脱氨基作用。

九、端粒酶

端粒酶(telomerase)是真核细胞染色体 DNA 复制所特有的也是必需的一种依赖于 RNA 的 DNA 聚合酶,它所起的作用是维持染色体端粒结构的完整。端粒(telomere)是位于一条染色体末端的特殊结构,由 DNA 和蛋白质组成,其中的 DNA 被称为端粒 DNA。端粒 DNA 由许多短重复序列组成,无编

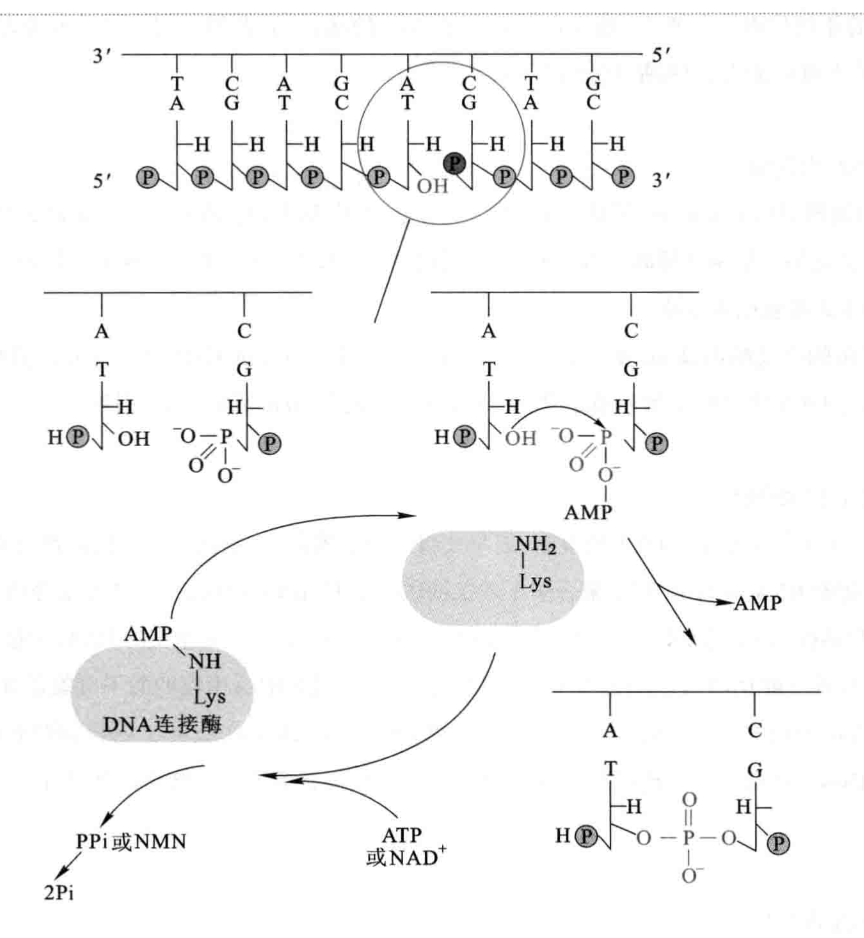


图 33-16　DNA 连接酶的作用机理

码功能。例如,人端粒 DNA 的重复序列是 TTAGGG,重复两千多次。端粒的主要功能是保护染色体,防止染色体融合、重组和降解。

当真核细胞的线形染色体 DNA 复制到一定阶段的时候,位于端粒 DNA3′端的冈崎片段上的 RNA 引物被切除,必然留下一段空隙。该空隙无法通过 DNA 聚合酶直接填补,原因是 DNA 聚合酶不能从 3′→5′方向催化 DNA 的合成。若上述空隙不及时填补,染色体 DNA 每复制一次,端粒 DNA 就会少一段。如何解决末端 RNA 引物切除后留下来的"烦恼"就依赖于端粒酶了。

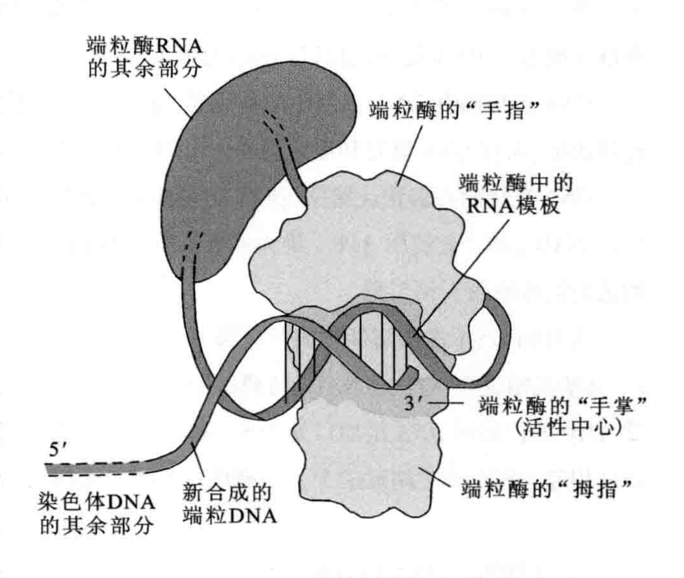
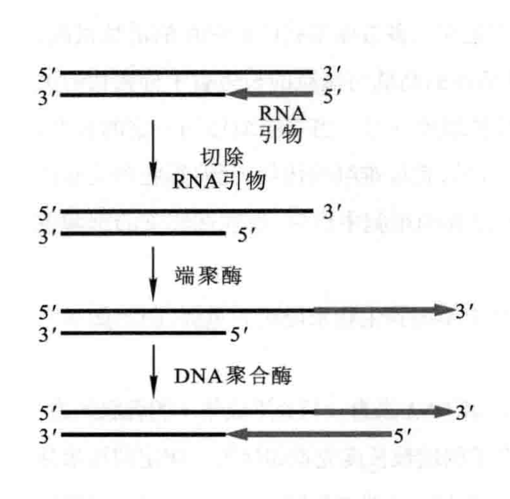


图 33-17　端粒酶的结构模型(Alberts 等,2007)

端粒酶也称为端聚酶,由蛋白质和 RNA 两种成分组成,其中蛋白质具有逆转录酶的活性,其三维结构与其他逆转录酶有明显的相似性,而 RNA 含有 1.5 拷贝的端粒 DNA 重复序列,这一部分序列作为逆转录酶的模板(图 33-17)。不同来源的端粒酶在 RNA 长度上变化很大,但它们都形成特定的二级结构,而作为模板的那一部分序列始终突出在外,处于单链状态,便于与端粒区的重复序列结合,并有效地充当逆转录的模板。

端粒酶的作用过程是(图 33-18):首先是其 RNA 中 0.5 拷贝的端粒 DNA 重复序列与端粒 DNA 最后一段重复序列配对,而剩余的 1 拷贝重复序列凸出在端粒的一侧作为模板。随后发生逆转录反应,



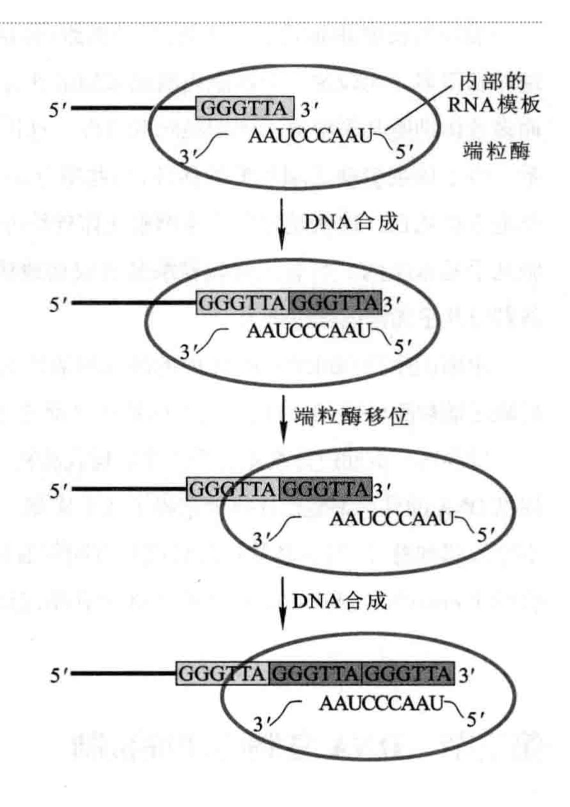


图 33-18　端粒酶的作用机理

在端粒 DNA 的 3′端填加 1 拷贝的重复序列。当逆转录反应结束以后，端粒酶移位，重复上面的反应，直到端粒突出的一端能够作为合成新的冈崎片段的模板，从而填补上一个冈崎片段 RNA 被切除后留下的空隙。

框 33-1　生化与健康——端粒激活与青春再现

近年来，有许多研究表明端粒缩短是衰老的主要原因。人类有一种早衰性遗传疾病叫先天性角化不良症（dyskeratosis congenita），这种疾病患者编码角化不良蛋白（dyskerin）的基因有缺陷，致使体内的端粒酶 RNA 的量比正常人低 5 倍，而使得患者体内所有细胞的端粒酶活性先天不足而引发早衰。而正常人体细胞内的端粒酶活性很低，这使得我们每一个人都会经历一个正常的衰老过程。有人通过某种方法激活端粒酶活性（telomerase activation，TA），则发现有助于延年益寿和抗衰老，而不会增加癌变或其他副作用的机会。例如，哈佛医学院的 Ronald A. DePinho 等人，在 2010 年 10 月 *Nature* 上发表的一篇论文就报道，他们从中草药黄芪中提炼的一种名为"TA-65"的有效成分，可以激活"端粒酶"。当"TA-65"被注射到老的小鼠体内后发现，小鼠的染色体端粒不仅延长了，衰老的脑细胞、免疫系统、脾细胞和生殖器官也恢复了活力，甚至连骨密度都提高了，近似于让年龄相当于 80 岁人的小鼠变得像幼鼠。再如，西班牙国立癌症研究中心的 Maria Blasco 等人，在 2011 年 4 月份的 *Aging Cell* 上的论文也报道，TA-65 能延长特别短的端粒，并在不同的器官中都有"复活"的功效，对改善健康跨度（health span）很有效。当他们在雌鼠的食物中添加 TA-65 后，发现了食用 TA-65 的雌鼠在许多健康跨度指标（葡萄糖耐受、骨质疏松和皮肤弹性等）上都有改善。另一份来自美国加利福尼亚州立大学的研究报告显示，人类老化的速率与平时的生活方式和遗传基因有着很大关系，那些有着稍短染色体端粒的人患心脏病的风险是带有长染色体端粒人群的 3 倍。而"TA-65"的存在将会让人类的每个细胞都活跃起来。

若是人类注射了这种"TA-65"，是否有相同的效果还有待观察，尽管现在市场上已有这种产品销售。如果它能在人体内，就像在试验动物身上一样，有相同的效果的话，那么"TA-65"为人类带来的不仅是美丽，更是健康……

端粒的长度并非固定不变的，生殖细胞（包括胚胎干细胞）和肿瘤细胞端粒的长度要比体细胞的端粒长得多。比较这三类细胞内的端粒酶活性，发现胚胎细胞和大多数癌细胞内的端粒酶活性很高，而多数体细胞几乎检测不到端粒酶的活性。这说明端粒酶活性的高低与端粒的长短有十分密切的关系。由于体细胞缺乏端粒酶的活性，因此每分裂一次，端粒就缩短一点。当端粒缩短到一定的长度，细胞必然死亡。这就是为什么体细胞在体外培养到几十代以后，就很难继续传代，而癌细胞和生殖细胞几乎是永生的。当有人将端粒酶基因成功地转染到体外培养的细胞中以后，发现被转染的细胞重新获得几乎无限的增殖能力。

Quiz10 四膜虫曾是纯化端聚酶最好的材料，为什么？

单细胞生物（如四膜虫）体内的端粒酶活性也很高，这对于单细胞生物来说极为重要，因为如果它们缺乏端粒酶活性，最终就必然走向物种灭绝之路。

世界第一例通过体细胞克隆产生的哺乳动物——多莉羊，其DNA来自一只6岁成年羊的乳腺细胞，提供DNA的乳腺细胞已在体外培养了几个星期。那么，多莉羊的端粒长度究竟如何呢？测定的结果并不令人感到意外，因为其端粒的长度只有同年通过正常生殖产生的羊端粒长度的80%。此外，威那综合征（Werner's syndrome）是一种遗传性早衰症，患者体细胞端粒被证实会比一般人以更快的速率变短。

第三节　DNA复制的详细机制

DNA复制是以复制子为单位进行的。任何一个复制子都含有一个复制起始区，有些复制子还含有特定的终止区。

不同基因组DNA具有不同的复制子结构，像细菌染色体、质粒、噬菌体、病毒和线粒体基因组DNA都只有一个复制子，而真核细胞内的每一个染色体DNA都含有多个复制子，许多古菌的单个染色体DNA也含有多个复制子。由于单复制子基因组DNA的复制相对比较简单，因此，目前很多单复制子基因组DNA复制的主要过程已为人所知。

一、以大肠杆菌为代表的细菌基因组DNA的"θ复制"系统

大肠杆菌基因组是一个共价闭环的DNA，只由一个复制子组成，其复制的详细过程目前已十分清楚，绝大多数与复制有关的蛋白质和酶已得到阐明（表33-4）。

► 表33-4　参与大肠杆菌DNA复制的主要蛋白质或酶的名称和功能

蛋白质名称	功能
DNA旋转酶	Ⅱ型拓扑异构酶，负责清除复制叉前进中的拓扑学障碍
SSB	单链结合蛋白
DnaA蛋白	复制起始蛋白，识别复制起始区 *oriC*
DnaB蛋白	DNA解链酶
DnaC蛋白	招募DnaB蛋白到复制叉
DnaG蛋白	DNA引发酶，引物合成
DnaT蛋白	辅助DnaC蛋白的作用
HU蛋白	类似于真核细胞的组蛋白，结合DNA并使DNA弯曲，有助于DnaA蛋白的作用
PriA蛋白	引发体的装配
PriB蛋白	引发体的装配
PriC蛋白	引发体的装配
DNA聚合酶Ⅲ	DNA链的延伸
DNA聚合酶Ⅰ	切除引物，填补空隙
DNA连接酶	缝合相邻的冈崎片段
DNA拓扑异构酶Ⅳ	分离子代DNA
Tus蛋白	复制终止

为了叙述上的方便，所有的 DNA 复制都可以分为起始、延伸、终止和分离三个阶段，现分别加以讨论。

(一) DNA 复制的起始

大肠杆菌的 DNA 复制起始于对 *oriC* 的识别(图 33-19)，结束于引发体的形成(图 33-20)。

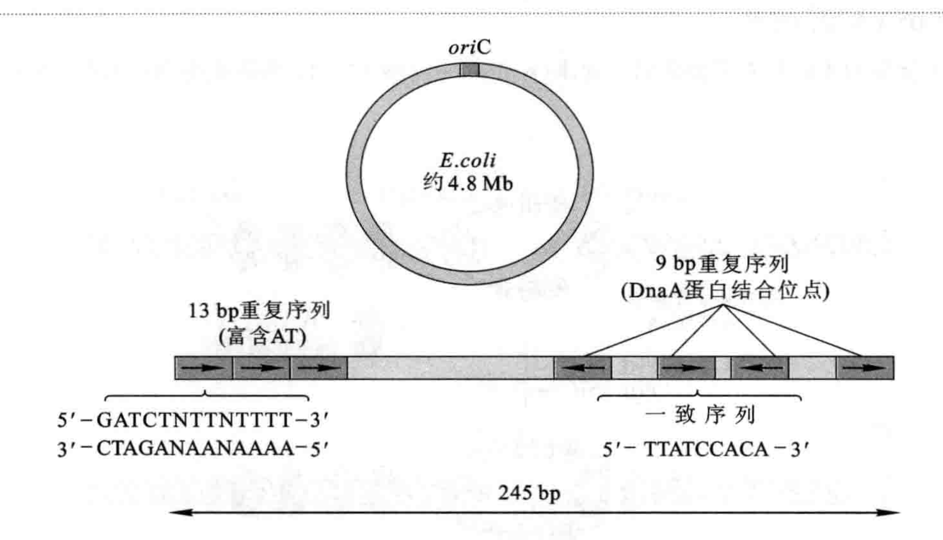


图33-19 大肠杆菌 DNA 复制起始区 *oriC* 的结构

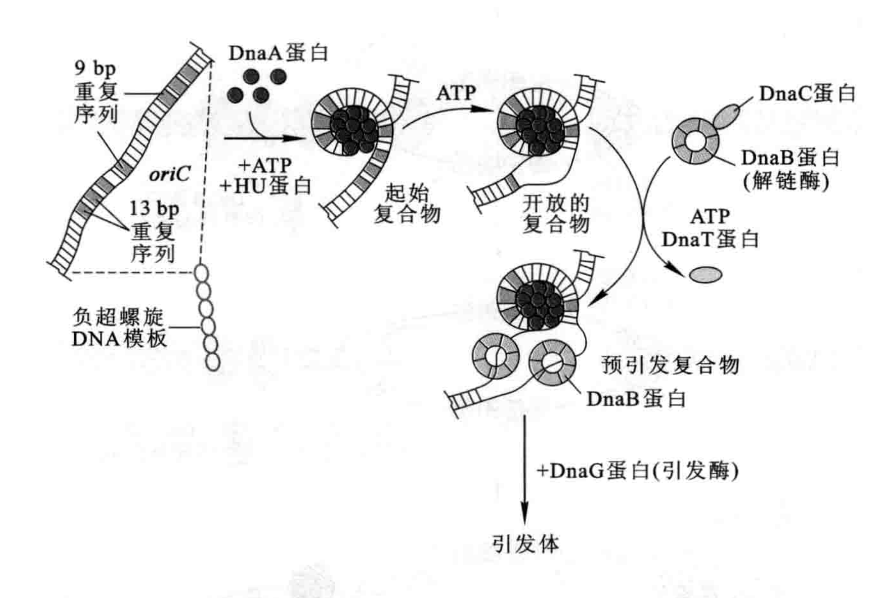


图33-20 大肠杆菌 DNA 复制过程中引发体的形成

整个起始阶段的主要反应依次是：

(1) 在两种 DNA 结合蛋白——HU 蛋白和整合宿主因子(integration host factor，IHF)的帮助下，带有 ATP 的 DnaA 蛋白四聚体识别并结合 *oriC* 的 9 bp 重复序列，这种结合具有正协同性。

(2) DnaA 蛋白相互之间聚合成蛋白核心，DNA 则环绕其上形成类似真核细胞的核小体结构。

(3) DnaA 蛋白水解结合的 ATP，以此驱动 13 bp 重复序列内富含 AT 的碱基对解链，从而形成长约 45 bp 的开放起始复合物。

(4) 在 DnaC/DnaT 蛋白的帮助下，两分子的解链酶(DnaB 蛋白)被招募到解链区，形成预引发体(preprimosome)。此过程需要消耗 ATP。

(5) 在 DnaB 蛋白的催化下，*oriC* 内的单链区域不断扩大，形成两个明显的复制叉。很快，SSB 开始与单链区结合。

(6) 两分子 DnaB 蛋白各自朝两个方向，沿着后随链的模板，催化两个复制叉的解链，DnaA 蛋白随

之被取代下来。

(7) 在 PriA、PriB 和 PriC 蛋白的帮助下，DnaG 蛋白被招募到复制叉，并与 DnaB 蛋白结合在一起，形成引发体。

(8) DnaG/DnaB 蛋白复合物沿着 DNA 模板链，先后为前导链和后随链合成 RNA 引物。

(二) DNA 复制的延伸

DNA 复制的延伸首先需要形成复制体(replisome)(图 33-21)，然后在引物的末端延伸前导链和后随链。

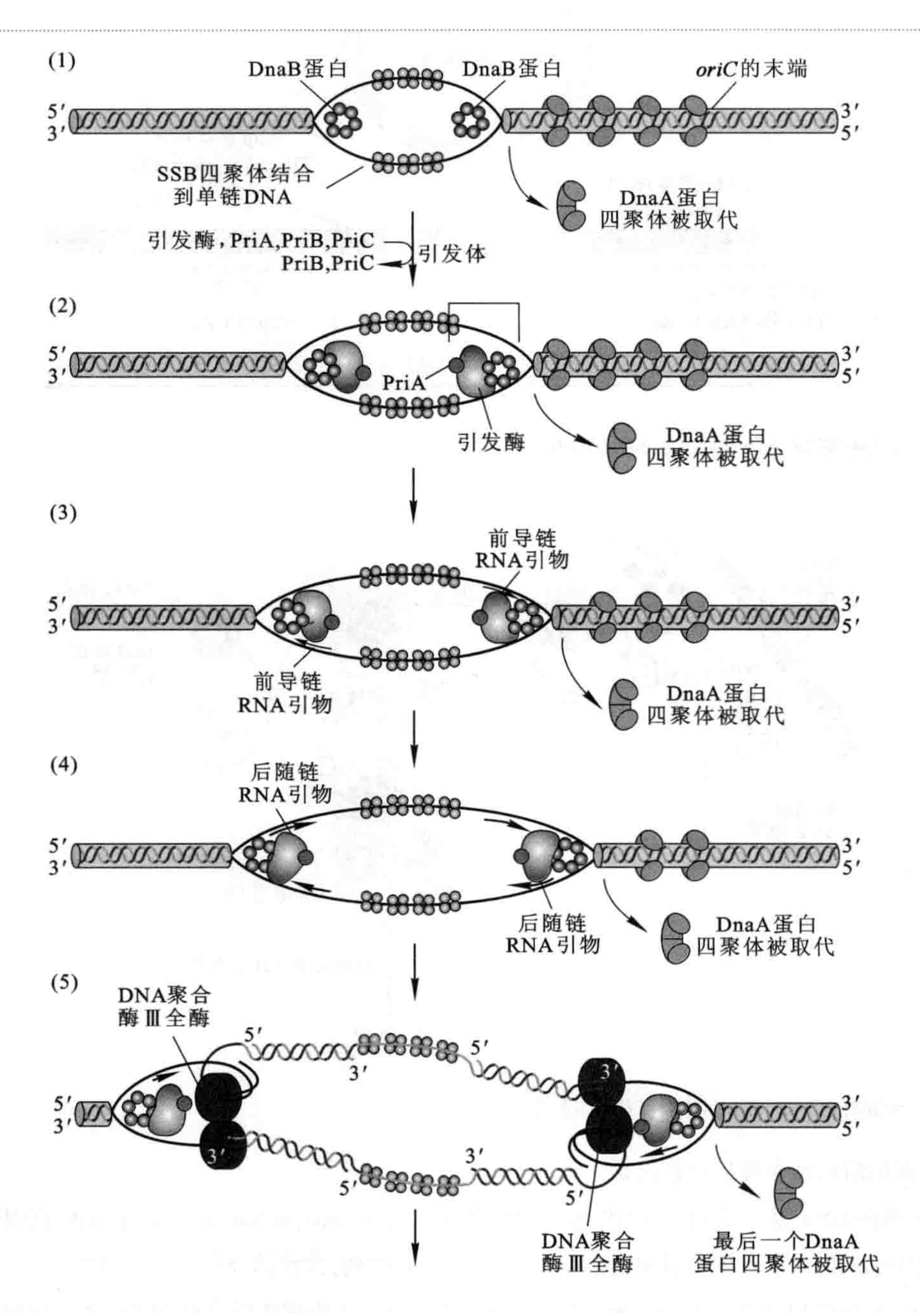


图 33-21　大肠杆菌 DNA 复制过程中复制体的形成(Voet 等，2011)

(1) 复制体的形成　当 DNA 聚合酶Ⅲ全酶加入到引发体上以后，真正意义上的 DNA 复制才开始。这种由 DNA 和多种蛋白质组装而成的催化 DNA 复制的复合体被称为复制体。对于一个进行双向复制的复制子来说，应该有两个复制体，每一个复制体能同时进行前导链和后随链的合成。

(2) 前导链合成　大肠杆菌的每一个复制体上都有一个由 2 个 DNA 聚合酶Ⅲ分子组成的不对称二聚体，分别负责前导链和后随链的合成。当一个复制叉内的第一个 RNA 引物被合成以后，其中的一个 DNA 聚合酶Ⅲ的全酶即可以在引物的 3′- 羟基上，连续地催化前导链的合成，直到复制的终点。

(3) 后随链合成　后随链的合成需要在合成每一个冈崎片段引物的时候，另外一个 DNA 聚合酶Ⅲ全酶的一部分暂时离开复制体。而在引物合成好以后，这个 DNA 聚合酶Ⅲ又需要重新装配，以启动 DNA 的合成。催化后随链合成的 DNA 聚合酶Ⅲ一旦遇到前一个冈崎片段的 5′端，即与 DNA 和滑动钳解离，但并没有离开复制体，而是与催化前导链合成的另一个 DNA 聚合酶Ⅲ结合在一起，以随时参与下一个冈崎片段的合成。留在冈崎片段上的滑动钳并没有立刻解体，而是依次将 DNA 聚合酶Ⅰ和连接酶招募到前一个冈崎片段的 RNA 引物处。被招募来的 DNA 聚合酶Ⅰ会及时地切除其中的 RNA 引物，并填补引物切除以后留下来的序列空白。与此同时，DNA 连接酶会将新的冈崎片段与前一个冈崎片段连接起来。随后，细胞内过量的处于游离状态的 δ 亚基像“扳手”一样，“撬开”活动钳，使其能够被循环利用。

总之，一轮冈崎片段合成的循环概括如下：DnaG 蛋白在复制叉与 DnaB 蛋白结合→ DnaG 蛋白启动后随链引物的合成→催化后随链合成的 DNA 聚合酶Ⅲ核心部分与 DnaG 蛋白相互作用，以限制引物的长度并暴露出引物的 3′-羟基端→钳载复合物 γ 亚基将环境中滑动钳 β 亚基装载到引物与模板的连接处→ DnaG 蛋白释放，DNA 聚合酶Ⅲ全酶合成新的冈崎片段，直至遇到上一个冈崎片段的 5′端→核心酶与 DNA 和滑动钳解离，但并没有离开复制体。

DNA 聚合酶Ⅲ复合物的装配与去装配由钳载复合物 γ 亚基控制。就大肠杆菌而言，一旦复制体装配完毕，γ 亚基仅在后随链上继续发挥作用。

(4) 前导链合成和后随链合成的协调　如图 33-22 所示，一个复制叉内的 DNA 聚合酶Ⅲ全酶二聚体同时催化前导链和后随链的合成，但它只能朝一个方向前进。酶之所以能够做到这一点，是因为后随链的模板在复制过程中形成突环结构，以便与前导链模板的方向保持一致。

(三) DNA 复制的终止和子代 DNA 的分离

大肠杆菌染色体 DNA 复制结束于终止区（terminus）（图 33-23），但终止区并不是 DNA 复制必需的，若是人为地将其去除，DNA 复制也能终止。在终止区的两侧存在两组由 23 bp 组成的终止子位点

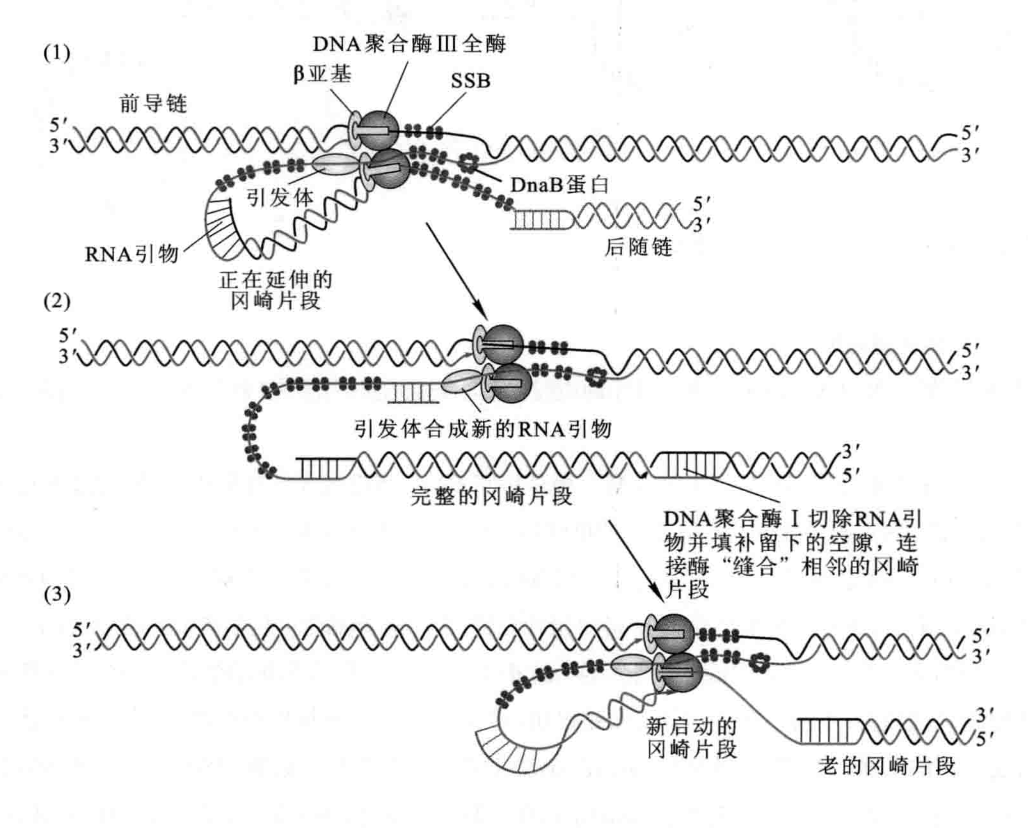


图 33-22　大肠杆菌 DNA 复制的延伸以及两条链合成的协调（Voet 等，2011）

(*Ter* 位点),左侧一组含有 *TerF*、*TerB* 和 *TerC*,右侧一组含有 *TerA*、*TerD* 和 *TerE*。*Ter* 位点富含 GT,有一种叫终止区利用物质的蛋白质(terminator utilization substance,Tus 蛋白)能特异性地结合这些位点。Tus 蛋白是解链酶 DnaB 蛋白的抑制剂,其分子表面有一个深深的带正电荷的裂缝,正好能与 15 bp 的 Ter 序列结合。当它与 *Ter* 位点结合以后可阻止复制叉的前进。因此,当位于终止区两侧的 *Ter* 位点上结合 Tus 蛋白以后,另一侧的复制叉便无法越过最后的 *Ter* 位点。

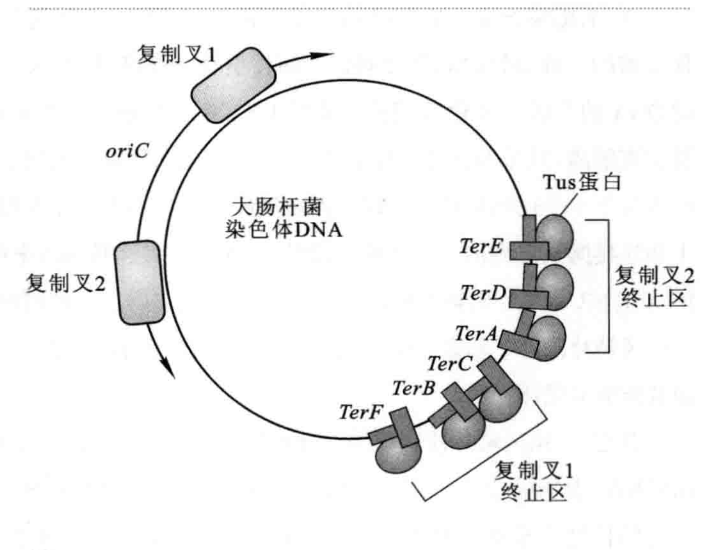


图 33-23 大肠杆菌 DNA 复制终止区的结构

当图中的复制叉 1 和复制叉 2 在终止区相遇后,DNA 复制即停止,那些位于终止区内尚未复制的序列(50~100 bp 长)会在两条母链分开以后,通过修复的方式填补。但无论如何,最后复制产生的两个子代 DNA 分子仍以连环体的形式存在,因此在分配给两个子细胞之前,必须进行"连体分离"手术。在大肠杆菌内,行使"主刀"任务的是拓扑异构酶Ⅳ或 XerCD 蛋白(图 33-24)。这两种蛋白质作为位点特异性重组酶,可识别终止区内的 *dif* 位点,切开 DNA 的两条链,在交换后进行再连接。

Quiz11 若是在大肠杆菌染色体 DNA 上人为插入第二个复制起始区,并与原来的复制起始区保持比较远的距离,你认为 DNA 复制还能正常地进行吗?

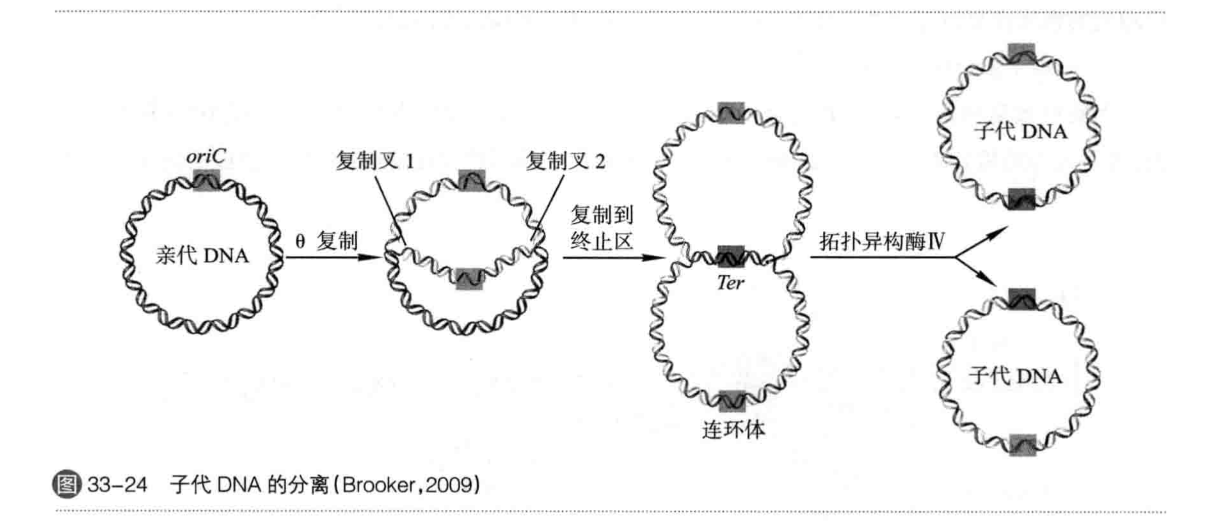


图 33-24 子代 DNA 的分离(Brooker,2009)

二、滚环复制系统

某些噬菌体(如 ΦX174)DNA 和一些小的质粒在宿主细胞内进行滚环复制(rolling-circle replication,RC 复制)。

以 M13 噬菌体为例,其基因组 DNA 是一种单链正 DNA,当进入大肠杆菌以后,即通过 θ 复制形成复制型双链 DNA,新合成的与正链互补的单链称为负链。复制型 DNA 形成以后,即进行滚环复制。滚环复制的全过程是(图 33-25):首先是位点特异性起始蛋白——A 蛋白识别并结合复制型 DNA 上的特殊位点;随后,具有内切酶活性的 A 蛋白切开正链,产生游离的 3′- 羟基和被共价固定的 5′- 磷酸(与 A 蛋白的一个 Tyr 残基以磷酸酯键相连);在宿主细胞 DNA 聚合酶Ⅲ的催化下,前导链(新的正链)开始合成,但它并不需要 RNA 引物,而是以切口处的 3′- 羟基直接作为引物,充当模板的是负链 DNA;随着新的正链被合成,负链环仿佛在滚动,老的正链则被取代后游离出双螺旋,以单链的形式存在;SSB 与游离出来的老的正链结合,新的正链在不断地合成,直到一条全新的正链得以复制;新的正链在合成好以后,仍然与老的正链以共价键相连,这时 A 蛋白可再次切开正链,使老的正链释放出

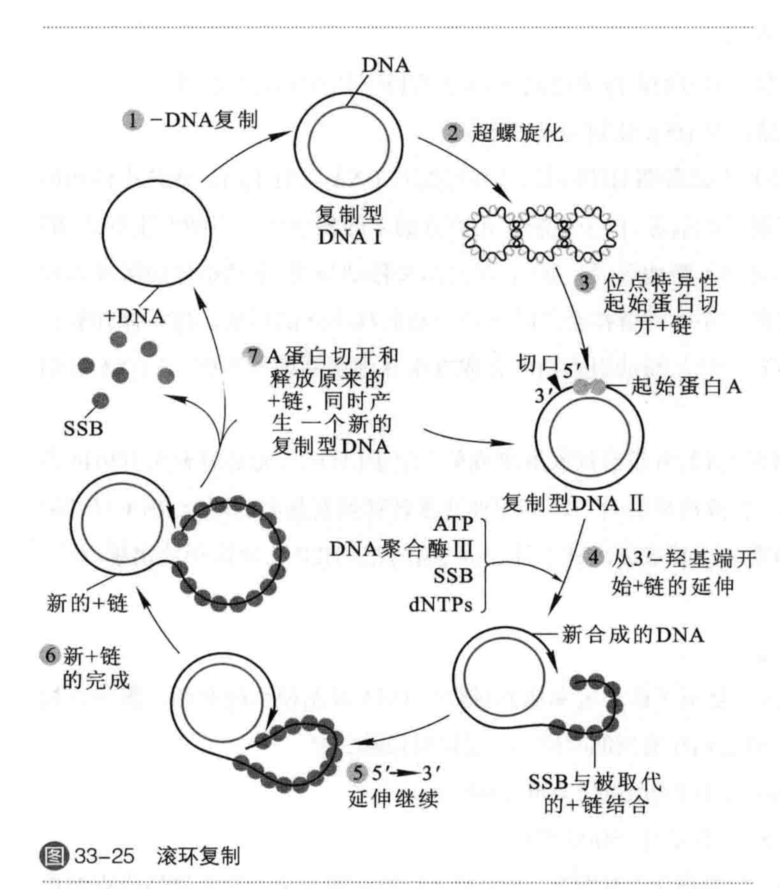


图 33-25　滚环复制

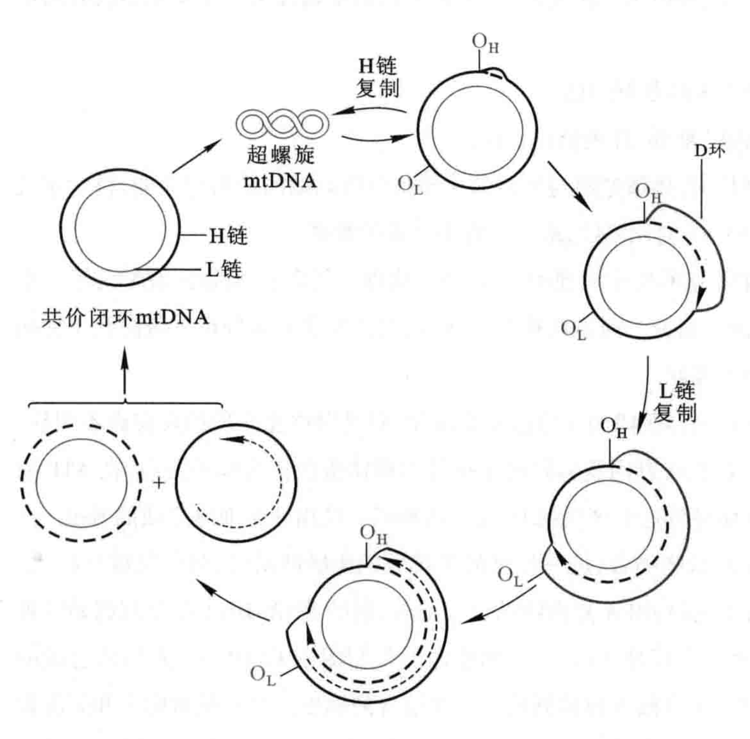


图 33-26　线粒体 DNA 的 D 环复制

来。负链则与新的正链形成新的复制型 DNA，并进行下一轮滚环复制。

滚环复制并不是少数噬菌体和质粒 DNA 所特有的复制模式。真核细胞的某些基因，例如两栖动物卵母细胞内的 rRNA 基因、哺乳动物细胞内的二氢叶酸还原酶基因，在特定的情况下会进行局部的滚环复制，以便在较短的时间内迅速增加目标基因的拷贝数。

三、D 环复制系统

线粒体和叶绿体是两种半自主性细胞器，它们自带的 DNA 通常以 D 环(displacement-loop，D 环)的方式进行复制。少数病毒，如腺病毒，也进行 D 环复制。

以动物细胞的线粒体 DNA (mtDNA)为例，组成它的两条链以共价闭环的形式存在，一条链因富含 G 而具有较高的密度，因此被称为重链(H 链)，另一条链因富含 C 而具有较低的密度，因而被称为轻链(L 链)。每一个 DNA 分子有两个复制起始区，一个用于 H 链的合成，称为 O_H，另一个用于 L 链的合成，称为 O_L。催化复制的酶是 DNA 聚合酶 γ，两条链的合成都需要先合成 RNA 引物，但都不形成冈崎片段，因此都是连续合成。

D 环复制的主要步骤(图 33-26)：① O_H 首先被启动，先合成 H 链；②新 H 链一边复制，一边取代原来的老 H 链。被取代的老 H 链以环的形式游离出来，这就是 D 环名称的由来；③当 H 链合成到约 2/3 的时候，O_L 被打开，由此启动 L 链的合成，新 L 链的合成是以被取代的 H 链作为模板，与 H 链合成的方向相反；④由于 L 链与 H 链合成在时间上的不同步，因此当新 H 链合成好的时候，L 链还有约 2/3 尚未复制；⑤先合成好的 H 链先被连接酶缝合，L 链则等合成好以后再进行连接。

四、真核细胞的细胞核 DNA 复制

真核细胞的核 DNA 复制在很多方面与细菌极为相似，但至少有以下几点不同于细菌：

(1) 需要解决核小体和染色质结构对 DNA 复制构成的障碍。

与细菌缺乏核小体和染色质结构形成鲜明对比的是，真核细胞核 DNA 与组蛋白形成核小体和染色质的结构。有证据表明，随着复制叉的前进，位于复制叉正前方的老的核小体会不断发生解体，同时在复制叉的正后方，新的核小体又在不断地形成。显然，在复制叉移动与新、老核小体的形成和解体的紧密偶联的过程中，每一个老的核小体的解体会伴随着两个新的核小体的形成。核小体在数目上的倍增需要有新的组蛋白的供应。幸好，新的组蛋白的合成发生在细胞周期的 S 期，与 DNA 复制同步。

在母链上一个老的核小体解体的时候，并没有彻底解离成单个组蛋白分子，而是解聚成 H3/H4 四聚体和 2 拷贝的 H2A/H2B 二聚体。释放出来的 H3/H4 四聚体直接转移到新复制的两个 DNA 双螺旋中的一个之上，再与其他的组蛋白重新组装成新的核小体。而老的 H2A/H2B 二聚体释放出来以后，便融入到新合成的组蛋白库中。

(2) 复制叉移动的速率远低于细菌。

(3) 具有多个复制子，这可以弥补复制叉移动速率偏低对整个 DNA 复制速率的制约。虽然真核细胞核 DNA 具有多个复制起始区，但它们在复制的时候不一定同时启动。

(4) 冈崎片段的长度为 100~200 nt，小于细菌的 1 000~2 000 nt。

(5) 复制被限制在细胞周期的 S 期，并受到严格的调控。

(6) 在第一轮复制结束之前不可能进行第二轮复制，而快速生长的细菌在第一轮复制还没有结束的时候就可以进行第二轮复制。

(7) 需要端粒酶解决染色体 DNA 末端复制问题。

(8) 复制终止没有特定的像大肠杆菌 *Ter* 序列的终止区。

为了深入了解真核细胞的复制机制，搞清楚参与复制的主要蛋白质和酶的结构与功能，科学家选择了一个简单适用的猿猴病毒 40（SV40）复制系统，取得了许多宝贵的数据。

SV40 最初是在野生猴子的肾细胞培养物中发现的一种 DNA 病毒。其染色体较小，复制几乎完全依赖于宿主蛋白，因此可以在体外进行研究。通过这种手段，研究人员发现有 8 种宿主蛋白和 1 种病毒自身编码的蛋白参与 SV40 的 DNA 复制。

目前对 SV40 DNA 复制的起始和延伸阶段的反应已非常清楚，只是对终止阶段的反应尚不明确，主要步骤包括（图 33-27）：① 2 分子由病毒基因组编码的 T 抗原六聚体蛋白作为起始蛋白，在 ATP 的存在下与 SV40 复制起始区结合，从而导致这个区域的 DNA 双链解链。②由 3 个亚基组成的 SSB——RPA 与上述稳定的起始蛋白 /DNA 复合物结合，进一步刺激 T 抗原的解链酶活性，致使解链区扩大。③ DNA 聚合酶 α/ 引发酶复合物与 T 抗原 /RPA 复合物结合。不久，新的 RNA-DNA 在复制起始区被合成，但 DNA 聚合酶 α 最多在引物之后延伸 15 nt。④细胞内的复制因子 C（RF-C）先与新合成的 DNA 3′端结合，然后将 PCNA 和 DNA 聚合酶 δ 招募到模板。⑤结合到模板上具有校对能力和高度进行性的 DNA 聚合酶 δ 取代 α，继续延伸新合成的 DNA 链。⑥前导链和后随链的合成分别经历了从 DNA 聚合酶 α/ 引发酶复合物到 DNA 聚合酶 ε 和 PCNA/DNA 聚合酶 δ 的转换。⑦前导链和后随链上的 RNA 引物都由 FEN1/RNaseH1 切除。由于 DNA 聚合酶 α 无自我校对活性，因此由它催化的约 15 nt 的片段更容易发生错配。幸运的是，FEN1 在切引物的最后一个核苷酸的时候，如果发现有错配的碱基，就可以发挥其内切酶的功能，顺便切掉包括错配碱基在内的 DNA 片段。⑧ DNA 聚合酶 δ 在催化下一个冈崎片段合成的时候，会将引物和错配的碱基切除后留下的空缺给补上。随后 DNA 连接酶 Ⅰ 将相邻的冈崎片段之间的缺口连上。⑨ DNA 拓扑异构酶 Ⅰ 负责清除复制叉移动中形成的正超螺旋，拓扑异构酶 Ⅱa 和 Ⅱb 则负责解连环体，促进最后 2 个以共价键相连的连环体 DNA 彼此分开。

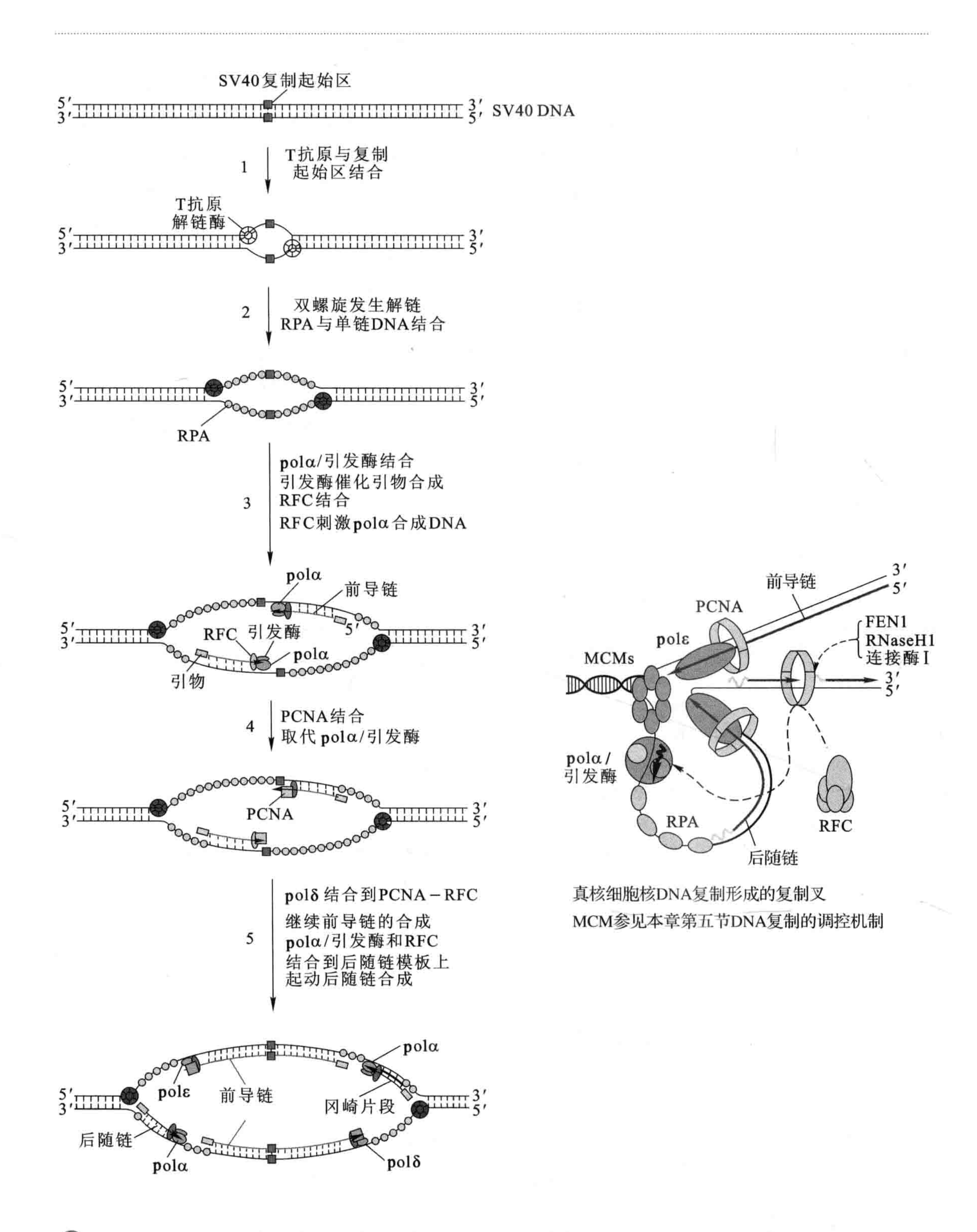


图 33-27　SV40 的 DNA 复制模型以及真核细胞 DNA 复制叉的结构模型（pol 为 DNA 聚合酶的缩写）

真核细胞核 DNA 复制的起始与 SV40 十分相似，包括启动复制所必需的蛋白质的结构与功能、导致复制启动的一系列事件发生次序，以及复制起始的调控机制等。然而，在复制起始点的选择上，却有一定程度的自由度。差别似乎集中在复制起始区的组成和起始蛋白识别的方式上。事实上，某些多细胞生物，如后生动物，在发育过程中能够改变复制起始区的数目和位置。

五、古菌的 DNA 复制

与细菌相似，古菌的染色体 DNA 也是单倍体的环状 DNA，其上的基因数目也差不多。然而，在 DNA 模板存在的状态、包装的方式以及复制机制上，古菌与真核生物非常相似。

所有生物体内的 DNA 在没有复制之前，绝大多数是以负超螺旋的状态存在的，但不同类型的生

物引入负超螺旋的机制并非相同。细菌是通过旋转酶引入，而真核生物利用组蛋白与 DNA 形成核小体的机制引入。许多古菌同时具有 DNA 旋转酶和组蛋白，因此，这些古菌在染色体 DNA 引入负超螺旋的手段可能类似于细菌，也可能类似于真核生物，还可能同时利用组蛋白和旋转酶。古菌的组蛋白长度要短于真核生物，但氨基酸序列及三维结构与真核生物相似，因此也可以自发组装成组蛋白核心。然而，古菌的组蛋白在与 DNA 形成核小体的时候，只形成四聚体核心，而不是真核生物八聚体核心。细菌没有组蛋白，虽然含有类似组蛋白的碱性蛋白与 DNA 结合，但从来不会形成核小体的结构。事实上，许多古菌也含有这样的碱性蛋白。此外，对于那些生活在极端高温下的古菌而言，它们含有 DNA 反旋转酶(reverse DNA gyrase)。与旋转酶不同的是，反旋转酶催化的反应是将正超螺旋引入染色体 DNA。这种反旋转酶的存在显然是为了维持这些嗜高温古菌细胞内 DNA 的稳定。

在 DNA 复制上，古菌与细菌相似的是，它也不需要也没有端聚酶，因为其染色体 DNA 是环状，无端粒结构，而在其他方面更像真核生物，主要表现在：① DNA 模板与组蛋白形成核小体，而核小体结构对复制会产生一定影响；②许多古菌具有多个复制起始区；③参与复制的许多蛋白质和酶在结构与功能上非常接近真核生物(表 33-5)。例如，参与 DNA 复制的主要聚合酶与真核生物一样，为 B 类，而细菌使用的是 C 类。

e33-2 DNA 聚合酶的分类

表 33-5 细菌、真核生物和古菌参与 DNA 复制的主要蛋白质和酶的比较

参与复制的蛋白质和酶	细菌	真核生物	古菌
起始蛋白	DnaA	ORC	Orc1/Cdc6
解链酶	DnaB	MCM 复合物	MCM
解链酶转载物	DnaC	Cdc6+Cdr1	Orc1/Cdc6
DNA 聚合酶	C 类	B 类	B 类
滑动钳	DNA 聚合酶Ⅲ的 β 亚基	PCNA	PCNA
滑动钳转载物	γ 亚基复合物	RF-C	RF-C
DNA 连接酶	依赖于 NAD^+	依赖于 ATP	依赖于 ATP
切除引物的酶	DNA 聚合酶Ⅰ/RNaseH	RNaseH/ FEN1	RNaseH/ FEN1
端聚酶	无	有	无

第四节 DNA 复制的高度忠实性

Quiz12 人类基因组 DNA (单倍体)每复制一次最多会有多少个错配的碱基?

DNA 复制具有其高度的忠实性(high-fidelity)，其错误率在 10^{-11}~10^{-9}，要远远低于 DNA 转录、RNA 复制和翻译的错误率(10^{-4}~10^{-3})。概括起来有以下几种机制在起作用：

(1) 4 种 dNTPs 浓度的平衡　细胞内 4 种 dNTPs 浓度的平衡对于 DNA 复制的忠实性有很大的影响，理想的比例应该是 1∶1∶1∶1。如果一种核苷酸的浓度远远高于另外 3 种，那浓度过高的核苷酸参入新合成 DNA 分子上的机会就大增，这会增加核苷酸错配的可能性而使复制忠实性下降。好在细胞内负责合成脱氧核苷酸的核苷酸还原酶具有非常精妙的调节机制，从而可以维持 4 种 dNTPs 浓度的平衡(详见第三十二章“核苷酸代谢”)。有一种在体外使用 PCR 对特定基因进行随机突变的方法，就是在反应系统中故意让其中一种 dNTP 的量很低而实现的。

(2) DNA 聚合酶的高度选择性　DNA 聚合酶直接催化核苷酸的参入，它能够根据模板链的核苷酸序列，按照碱基互补配对的规则选择正确的核苷酸。研究表明，酶使用了几何选择(geometric selection)和构象变化两种机制。若是正确的核苷酸进入酶的活性中心，那形成的碱基对在外形、距离和糖苷键连接的角度上匹配得几乎完美；若是错误的核苷酸进入活性中心，形成的碱基对在外形、距离和糖苷键连接的角度上与 Watson-Crick 碱基对就有较大的差别。于是错误的 dNTP 因几何形状不对而不适合进入活性中心，相反，正确的 dNTP 很容易进入酶的活性中心。在没有底物结合的时候，聚

合酶的手指－手掌－拇指的构象处于开放状态。一旦正确的 dNTP 进入酶活性中心，“诱导契合”开始起作用。正确配对的核苷酸会诱导酶的构象迅速发生较大的变化。构象发生变化的酶能更好地包被碱基对，使之采取合适的取向。如果错误的核苷酸进入活性中心，酶构象发生的变化就要慢近一万倍！

(3) DNA 聚合酶的自我校对　这是由 DNA 聚合酶本身所具有的 3′－外切酶活性来执行的（详见“DNA 聚合酶的结构与功能”）。

(4) 错配修复　以上三种机制并不能保证合成好的新链绝对不会出现任何错配的核苷酸。倘若真的出现了错配的核苷酸，机体还有没有什么补救的措施？实际上，在细胞里早已预备好最后一道“防线”专门来处理复制中错配的核苷酸，这种机制称为错配修复（mismatch repair），详见下一章“DNA 损伤、修复和突变”。

(5) 使用 RNA 作为引物　在 DNA 复制中要先合成 RNA 引物曾叫人百思不得其解。使用 RNA 引物似乎既浪费能量，又浪费时间。那么，合成 RNA 引物的真正意义何在？如果从复制的忠实性这个角度去理解，就找到了答案。在 DNA 合成刚刚开始的时候，最初参入的核苷酸难以与模板链形成稳定的双螺旋，因此很容易发生错配，但如果使用 RNA 引物，就可以最终清除这些错配的核苷酸，因为 RNA 引物迟早是要被水解的。

第五节　DNA 复制的调节机制

DNA 复制的调控主要集中在起始阶段，现分别介绍细菌和真核细胞的 DNA 复制的起始过程是如何受到调控的。

一、细菌 DNA 复制起始的调控

以大肠杆菌为例，何时进行新一轮 DNA 复制由 DNA 腺嘌呤甲基化酶（DNA adenine methylase，Dam）和 DnaA 蛋白控制。

如图 33-28 所示，亲代 DNA 分子的两条链均被甲基化，甲基化位点是 5′-GATC-3′序列中腺嘌呤的 6 号位 N 原子，催化甲基化的酶是 Dam，甲基供体是 SAM。*oriC* 内含有 11 个重复的 GATC 序列，在亲代 DNA 分子上都被甲基化了。只有甲基化的 *oriC* 才能被复制起始蛋白（DnaA 蛋白）识别和结合，从而启动 DNA 的复制。

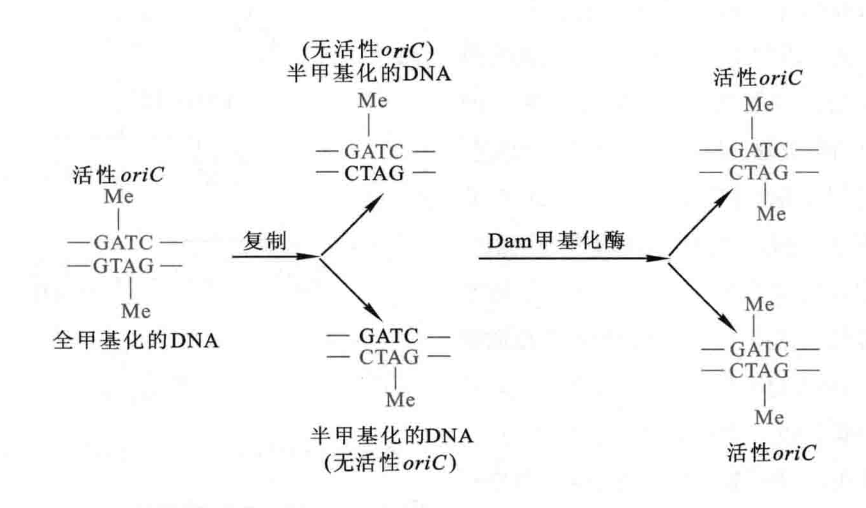


图 33-28　甲基化对细菌 DNA 复制的调节（Me 表示甲基）

Quiz13 如果让编码 SeqA 蛋白的基因缺失，那么大肠杆菌 DNA 复制的起始有何变化？

然而，新合成的 DNA 链还没有被甲基化，这意味着刚刚形成的两个子代 DNA 分子中只有一条链是甲基化的（DNA 的半保留复制）。在这种情况下，与细胞膜结合的一种名为 SeqA 的抑制蛋白与 *oriC* 结合，使得 DnaA 蛋白不能识别和结合 *oriC*。当 Dam 将子代 DNA 分子上的 GATC 序列甲基化以后，SeqA 立刻与 *oriC* 解离，DnaA 蛋白便开始与 *oriC* 结合，从而启动新一轮 DNA 的复制。

二、真核细胞 DNA 复制起始的调控

真核细胞的基因组 DNA 由多个复制子组成，各复制子的复制被限制在细胞周期的 S 期，且在一个细胞周期内，每一个复制子的复制只会启动一次。那么，真核细胞是如何做到这一点的呢？

最新的研究表明，真核细胞控制 DNA 复制启动的机制主要是由执照因子（licensing factor）控制的正调控。以酵母 DNA 复制的启动为例，首先需要在各复制起始区 ARS 上形成由起始区识别蛋白组装而成的复合体（origin recognition complex of protein，ORC）。ORC 由 6 个不同的亚基组成（Orc1~Orc6），存在于 DNA 复制的全过程。于是，ORC 和 DNA 之间的相互作用决定了复制的起点。但 ORC 并不能让 DNA 发生解链，因此虽然它在细胞周期的 G_1 期已经形成，但在此时复制并不能启动。复制何时启动是由另外一类称为执照因子的蛋白质控制。执照因子在细胞周期的 G_1 期合成，并转移到细胞核，在被招募到 ORC 上以后，开始刺激 DNA 复制的启动，这时细胞周期已进入 S 期。

执照因子包括 Cdc6（cell division cycle 6）蛋白、Cdt1 蛋白和 MCM（mini-chromosome maintenance）

框 33-2　生化研究动态——干细胞分裂过程中的染色体分离

干细胞是具有自我更新、高度增殖和多向分化潜能的细胞群体。这些细胞可以通过细胞分裂来维持自身细胞群的大小，同时又可以进一步分化成为各种不同类型的组织细胞，从而构成机体各种复杂的组织和器官。由于干细胞具高度分裂性，如小鼠的小肠黏膜干细胞在小鼠的一生中要分裂大约 1 000 次，而每分裂一次，DNA 就要复制一次，其中复制出错是不可避免的。如果细胞分裂过程中 DNA 的分配是随机的话，在干细胞内就会积累越来越多的错误（从干细胞分化出来的体细胞虽然也继承一些复制的错误，但其有限的生命周期不会造成错误的不断积累），这势必影响到干细胞基因组 DNA 的稳定性，而大大增加细胞癌变的危险。那么，干细胞是如何在不断的分裂过程中维持其基因组 DNA 的高度稳定性的呢？

在 1966 年前后，对小鼠胚胎培养细胞和中国仓鼠卵巢培养细胞的实验显示，老的 DNA 链和新的 DNA 链在子细胞中的分配并不是随机的：新的链倾向于缠绕在一起。1975 年，John Cairns 在其对上皮细胞分裂研究的基础上，提出"不朽链假说"（immortal-strand hypothesis）来解释上述现象。

该学说的主要内容是（图 33-29），干细胞的分裂是不对称的，而分裂的不对称性使以后的干细胞能够选择性得到含有最老的 DNA 模板的染色体。这样的假设是合乎逻辑的，因为如果一个干细胞在分裂中抓住最老的 DNA，就等于将复制可能产生的错误"拒之门外"，从而大大降低了细胞癌变的可能性，而将来分化成体细胞的子细胞得到易错的新 DNA 也没有太大危险，原因是它们很快停止分裂或者死亡，特别是在高度更新的组织。

迄今为止，"不朽链假说"已在很多类型的干细胞分裂系统中得到验证。

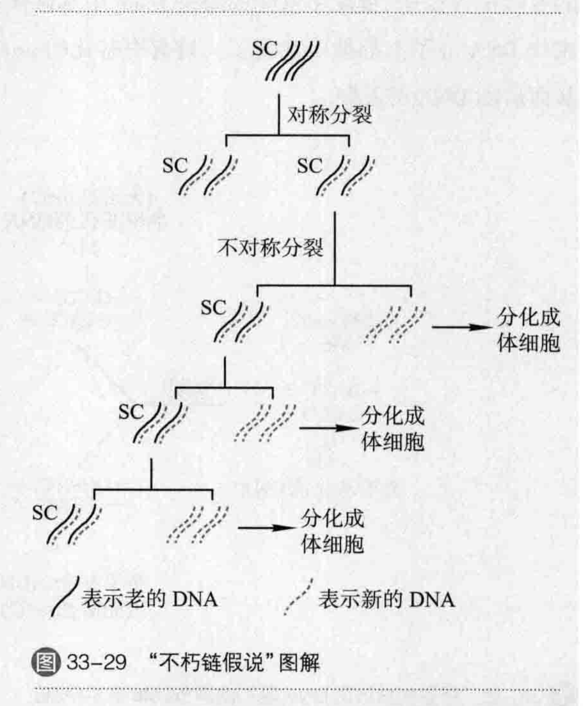


图 33-29　"不朽链假说"图解

蛋白。其中 Cdc6 蛋白和 Cdt1 蛋白先与 ORC 结合，然后再将 MCM 蛋白装载到 ORC 上。只有被 MCM 包被的 DNA 才能复制。具体过程是(图 33-30)：在 Noc3 蛋白的帮助下，Cdc6 蛋白与 ORC 结合后，允许 Cdt1 装载 MCM(Mcm2~Mcm7)上来，形成预复制复合物。Noc3 是一种与复制起始区相连的高度保守的 DNA 结合蛋白，与 ORC 和 MCM 蛋白都有相互作用。Mcm2 和其他 5 种 MCM(Mcm3~Mcm7)以六聚体的形式存在，具有解链酶的活性，负责亲代 DNA 的解链。此过程受到 RPA 的促进。DNA 合成受到 Mcm10 加入的引发，但 Mcm10 只有在受到依赖于 Dbf4 的蛋白质激酶(Dbf4-dependent protein kinase，DDK)——Cdc7/Dbf4 或依赖于周期蛋白的蛋白质激酶(CDK2/ 周期蛋白 E 和 CDK2/ 周期蛋白 A)的作用发生磷酸化修饰后才有活性。磷酸化的 Mcm10 招募 Cdc45、Sld 和 GIN 蛋白“护送”DNA 聚合酶 α / 引发酶到预复制复合物，在 ORC 或附近启动 RNA 引物的合成，由此引发 DNA 的复制。

DNA 复制一旦启动，Cdc6 蛋白和 Cdt1 蛋白就离开 ORC，其中 Cdt1 被泛酰化后进入蛋白酶体降解，MCM 蛋白则随着复制叉的前进逐步与 DNA 解离。

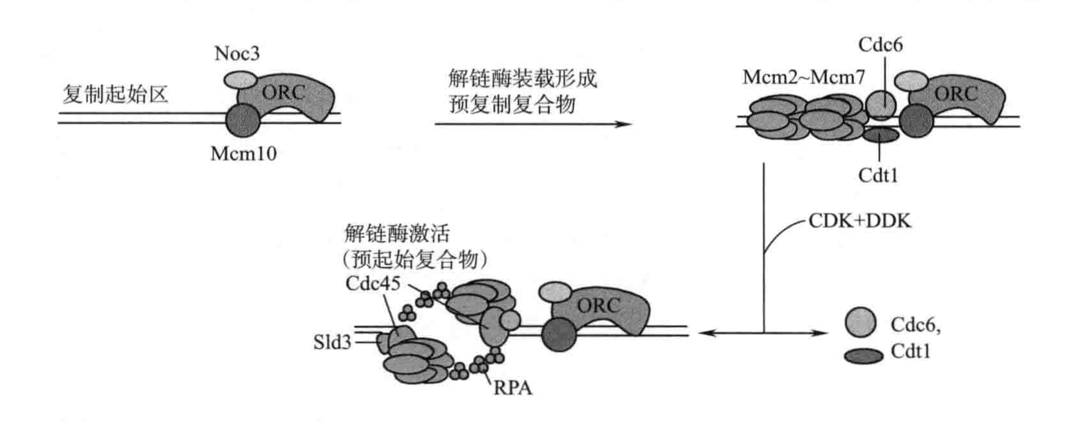


图 33-30 酵母细胞 DNA 复制的正调控

科学故事——半保留复制的实验证明

Watson 和 Crick 于 1953 年 4 月 25 日在 *Nature* 杂志上发表了具有划时代意义的题为“Molecular Structure of Nucleic Acids”的论文最后一段这样写道：“It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material. The structure itself suggested that each strand could separate and act as a template for a new strand，therefore doubling the amount of DNA，yet keeping the genetic information，in the form of the original sequence，intact.” 翻译成中文的意思是，“特定的碱基配对性质让我们立刻注意到遗传物质可能的复制机制。结构(双螺旋)的本身意味着每一条链可以分离开来，作为新链的模板，于是 DNA 的量加倍了，而且保证了遗传物质以原来原封不动的序列形式存在。”在文中，他们虽然对 DNA 复制的具体机制并没有提及，但已经就 DNA 复制采取的可能是一种半保留方式做出了大胆的预测。然而，细胞内 DNA 复制是否就是以这种方式进行的，光凭预测是不够的，还必须用实验去验证。由 Meselson 和 Stahl 设计的被誉为“生物学最美丽的实验”最终证明了 Watson 和 Crick 预测的正确性。



图 33-31 Meselson(左)与 Stalh 在最初相遇的一棵树下 42 年后重逢时的合影

1953 年，Meselson 在加利福尼亚州立理工学院(Caltech)开始了以化学为专业的研究生阶段的学习。他进入了 Pauling 的实验室，最终完成了其博士论文第二部分，关于 *N*,*N*- 二甲基丙二酰胺(*N*,*N*-dimethyl malonamide)晶体结构的工作。这部分工作的重点是确定这个分子中含有的肽键是不是共平面的，即是否与 Pauling 的共振理论相符。与他的博士论文第一部分，即关于大分子的密度梯度平衡沉降技术在 DNA 研究上的应用相比，第二部分鲜为人知。身为 Pauling 的弟子，在听过 Pauling 开设的有关化学键的课程以后，他对当 H 被其同位素氘(^2H)取代以后氢键的相对强度产生了

兴趣。就在他对氘参入到生物分子以后,生物体如何能够生存下来的问题发生兴趣的时刻,Meselson 正好去听了 Jacques Monod 在 Caltech 作的有关细菌诱导酶合成的学术报告。Meselson 推测,如果细菌能在重水里培养,然后被转移到一般的水里继续培养,则一旦加入诱导物,则任何新合成的蛋白质就应该具有正常的密度,与原来老的蛋白质的密度是有差别的,利用密度差异可以将"老"的蛋白质和新合成的蛋白质分开。事实上,如果在一个具有中间密度的溶液中离心,"老"的蛋白质可能沉降下来,而新合成的蛋白应该悬浮在上面。

不久,他开始将注意力转移到 DNA 复制的问题。在与 Max Delbrück 对 DNA 双螺旋结构以及可能的复制方式进行了一番热烈的讨论以后,他突然想到相同的方法也许可以用来研究 DNA 复制。于是,他决定投入精力去确定 DNA 复制是否就是按照 Watson 和 Crick 预测的半保留方式进行的。与此决定无关的一件事也许对其最后的成功带来决定性的影响:1954 年的夏天,Meselson 去马里兰州位于 Woods Hole 的海洋生物实验室,协助 Watson 进行一些滴定实验,以获得 RNA 双螺旋结构的证据。巧合的是,来自 Rochester 大学的 Frank Stahl 博士也在 Woods Hole 修生理学课程。一天,正当 Stahl 在树下思考噬菌体遗传性的一个问题的时候,他们相遇并从此成了朋友和学术伙伴。就在 Meselson 对噬菌体遗传学还似懂非懂的时候,他在微积分上的专长帮助 Stahl 解决了问题。不久 Meselson 提出了利用 Stahl 在研究噬菌体 DNA 上的特长,使用噬菌体 DNA 来合作研究 DNA 复制的可能性。他还提出了使用氘来重标记 DNA 的设想,以此用密度梯度离心来分离重的 DNA 和轻的 DNA。然而,他们很快就意识到使用噬菌体 DNA 进行实验带来的复杂性问题,因为噬菌体 DNA 之间存在高频的重组事件,由此引起亲代和子代 DNA 片段的交换是可以预见的,这势必会影响到结果的分析。幸运的是,Stahl 已经打算去 Caltech 做博士后研究,以方便以后的讨论和合作。既然噬菌体行不通,较为简单的单细胞系统——*E. coli* 也许是一个不错的选择。后来,Meselson 想知道更多有关 DNA 合成前体分子性质的知识,在文献调研中,得知胸腺嘧啶的类似物——5- 溴尿嘧啶(5-bromouracil,5-BrU)能够代替 T 参入到正在合成的 DNA 分子之中。5-BrU 相当于 T,由 Br 原子代替嘧啶环 C5 上的甲基。Br 与甲基具有几乎相同的范德华半径,但由于 5-BrU 与 T 的离子化程度不同,Meselson 想到也许可以用 5-BrU 标记 DNA,然后利用电泳的方法将它与含有 T 的 DNA 分离开来。特别重要的是,他特别想到含有 5-BrU 的 DNA 要比含有 T 的 DNA 重得多!于是,他考虑使用 5-BrU 来重标记 DNA 以验证 DNA 半保留复制的可能性。由于需要超离心技术的关系,他认识了 Caltech 的超离心技术奇才 Jerry Vinograd,并学会了如何操作当时最先进的 Beckman Spinco E 型超离心机。在 Vinograd 悉心指导下,Meselson 开始尝试用 7 mol/L 的 CsCl 重盐溶液来沉降 DNA。使用 CsCl 的主意来自他认为可以将重标记的 DNA 与轻 DNA 分开来的设想——重标记 DNA 沉降下来,轻 DNA 悬浮在表面。然而,令他们惊诧的是,在高速离心场下,一种盐密度梯度很快就形成了,且 DNA 迁移到与其等密度的区域,形成很窄的条带。于是有关密度梯度离心的概念被 Meselson 提出来,相关的论文发表在 1957 年 5 月份的美国科学院院刊上。论文中的图和理论计算实际上是他博士论文第一部分的内容。论文中还记录了含有 5-BrU 的 DNA(在含有 5-BrU 培养基中培养的受 T4 噬菌体感染 *E. coli* 中,得到被 5-BrU 标记的 T4 噬菌体 DNA)在密度梯度离心中的沉降情况。结果显示,被 5-BrU 标记的 DNA 条带所处位置的密度是 1.8 g/cm^3,而含有 T 的 DNA 所在位置的密度是 1.7 g/cm^3。虽然文中没有提到使用这种方法研究 DNA 复制,但用于研究完整病毒和生物大分子的报道还是非常富有开创性的。

幸运之神似乎正在一步一步地将 Meselson 和 Stahl 带到研究 DNA 复制的里程碑的实验设计思路上去。他们本来想用 5-BrU 去重标记 DNA 的,但后来担心 5-BrU 能诱发 DNA 突变而对细胞产生毒性,以及能否获得标记的均一性的问题,他们决定使用 $^{15}NH_4Cl$ 作为唯一 N 源的合成培养基,去培养 *E. coli* 以获得被 ^{15}N 标记的重 DNA。他们首先将 *E. coli* 放在 $^{15}NH_4Cl$ 培养基中连续培养十几代,得到了几乎全被 ^{15}N 标记的 *E. coli* DNA;然后,随着细菌的指数生长,用 10 倍过量的 $^{14}NH_4Cl$ 稀

释培养基。期间在不同的时段，将 DNA 从细菌中抽取出来，并使用 CsCl 密度梯度离心的方法进行分析。

研究的结果是，起初完全被 ^{15}N 标记的 DNA 具有单一的条带，而在 $^{14}NH_4Cl$ 培养基中培养的第一代细菌 DNA 的密度介于 ^{15}N-DNA 和 ^{14}N-DNA 之间，为杂交带 DNA（$^{15}N/^{14}N$-DNA）。随着杂交带的出现，亲代的条带（^{15}N-DNA）消失了。而在第二代细菌中，得到几乎等量的 ^{14}N-DNA 和杂交带 DNA。随着培养代数的增加，杂交带始终存在，而且它的量维持不变，但 ^{14}N-DNA 的量越来越多。由此可以得出结论，DNA 复制以半保留的方式进行。为了进一步确认这一点，Meselson 和 Stahl 在密度梯度离心之前，用热变性处理 CsCl 溶液中 DNA 样品（100℃下 30 min），结果发现，杂交 DNA 分成了 ^{15}N-DNA 和 ^{14}N-DNA 两条带。

e33-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/DNA_replication（维基百科有关 DNA 复制的内容）
2. http://www.wiley.com/college/pratt/0471393878/student/animations/dna_replication/index.html（内有 DNA 复制过程的动画）
3. http://themedicalbiochemistrypage.org/dna.html（完全免费的医学生物化学课程网站有关 DNA 的内容）
4. http://www.hhmi.org/biointeractive/dna/animations.html（内有 DNA 复制、转录和翻译等过程的动画显示链接）
5. http://web.virginia.edu/Heidi/chapter29/chp29frameset.htm（美国弗吉尼亚大学大学生物化学在线课程有关 DNA 复制的内容）

参考文献：

1. Koutroumpas K, Lygeros J. Modeling and analysis of DNA replication. *Automatica*, 47:1 156–1 164, 2011.
2. Pavlov Y I, Shcherbakova P V. DNA polymerases at the eukaryotic fork–20 years later. *Mutation Research*, 685:45–53, 2010.
3. Remus, Diffley J F X. Eukaryotic DNA replication control: Lock and load, then fire. *Current Opinion in Cell Biology*, 21:771–777, 2009.
4. Hubscher U, Maga G, Spadari S. Eukaryotic DNA polymerases. *Annu Rev Biochem*, 71:133–63, 2002.
5. Nishitani H, Lygerou Z. Control of DNA replication licensing in a cell cycle. *Genes Cells*, 7:523–534, 2002.
6. Davey M J, O'Donnell M. Mechanisms of DNA replication. *Curr Opin Chem Biol*, 4:581–586, 2000.
7. Lingner, Cech T R. Telomerase and chromosome end maintenance. *Curr Opin Genet Dev*, 8:226–232, 1998.
8. Zakian V A. Telomeres: beginning to understand the end. *Science*, 270:1 601–1 607, 1995.

第三十四章　DNA 的损伤、修复和突变

DNA 与其他生物大分子一样，在各种内外因素的作用下，其结构可能会遭受到损伤。但和其他生物大分子不一样的是，DNA 是唯一一种在发生损伤以后可以被完全修复的分子，而其他生物大分子在受到损伤以后要么被降解，要么被取代。当然，并不是所有发生在 DNA 分子上的损伤都可以修复。如果 DNA 受到的损伤不能及时修复，不仅会使 DNA 的复制和转录受到影响，还可能导致机体发生突变。

细胞在 DNA 受到损伤以后，之所以选择的处理方法是尽量将其修复而不是降解，是因为作为遗传物质的 DNA 分子在细胞内只有一个拷贝，如果将其降解的话，细胞也就失去了存在的根基，此外，DNA 的互补双螺旋结构使修复一个受损伤的 DNA 分子变得相对容易。正因为如此，一种生物体，即使是那些基因组甚小的生物，也会在 DNA 修复上投入大量的基因，这再次说明了"遗传物质的稳定性压倒一切"。

本章将集中介绍导致 DNA 损伤的各种因素、DNA 损伤的类型、各种修复机制和 DNA 的突变。

第一节　DNA 的损伤

导致 DNA 损伤的根源有细胞内在因素和环境因素，属于内在因素的有：① DNA 复制中发生的错误；② DNA 结构本身的不稳定，如 C 自发脱氨基生成 U，A 自发脱氨基生成 I；③细胞内活性氧（ROS）的破坏作用。属于环境因素的有物理因素和化学因素，其中前者包括紫外辐射和离子辐射（X 射线和 γ 射线），后者包括各种化学诱变剂，有的是天然的，如黄曲霉素（aflatoxin），有的是人造的，如顺铂（cisplatin）、芥子气和烷基化试剂等。

不同的因素通常引起不一样的损伤，一般根据受损的部位，将 DNA 损伤分为碱基损伤和 DNA 链的损伤两大类。

碱基损伤又可以分为 5 个亚类：①碱基脱落（base loss），损伤的原因是由于 DNA 结构的不稳定，更具体地说是 DNA 分子上连接碱基和核糖的糖苷键自发地发生了水解，以脱嘌呤最普遍；当细胞受热或酸度提高时，这种损伤加剧；②碱基转换，即由一种碱基转换为另外一种碱基；损伤的原因也是由于 DNA 结构的不稳定，但直接的原因是含有氨基的碱基自发地发生了脱氨基反应（亚硝酸的存在会加剧这种损伤），例如，C 和 A 经脱氨基反应分别转变为 U 和 I；③碱基修饰，这是某些化学试剂或活性氧直接作用于碱基造成的，例如活性氧修饰鸟嘌呤产生 8- 氧鸟嘌呤（8-oxoguanine）（图 34-14），烷基化试剂修饰鸟嘌呤产生 6- 甲基鸟嘌呤；④碱基交联，紫外辐射可导致 DNA 链上相邻的嘧啶碱基，主要是 T 之间形成嘧啶二聚体或 6-4 光产物（图 34-1）；⑤碱基错配，在 DNA 复制过程中，4 种 dNTP 浓度的不平衡、碱基的互变异构或碱基之间的差别不足都可能引起错配；虽然绝大多数错配碱基能被 DNA 聚合酶的自我校对机制切除，但仍会有少数"漏网之鱼"得以逃脱。

DNA 链的损伤又分为 4 个亚类：①核糖核苷酸在复制中参入到 DNA 链上；② DNA 链的断裂，有单链断裂和双链断链，原因包括离子辐射和某些化学试剂的作用，如博来霉素（bleomycin），相对于其他类型，这种损伤最为严重；当 DNA 出现太多的裂口，尤其是双链裂口的时候，往往难以修复，这会导致细胞的死亡，肿瘤放疗就基于此原理；③ DNA 链的交联，原因主要是某些双功能试剂的作用，导致 DNA 发生链间交联，如丝裂霉素 C（mitomycin C）和顺铂；④ DNA 与蛋白质之间的交联，UV 或甲醛可

图 34-1 紫外线引起的碱基损伤

诱导 DNA 结合蛋白与 DNA 之间形成共价交联。

第二节 DNA 的修复

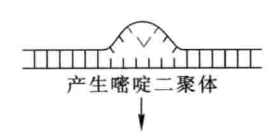
DNA 损伤的形式虽然很多，但细胞内的修复系统也不少，可谓是“魔高一尺，道高一丈”。基本上每一种损伤在细胞内都有相应的修复系统，有时还不止一种。根据修复的原理，DNA 修复可分为直接修复(direct repair)、切除修复(excision repair)、双链断裂修复和损伤跨越等几类，下面就这几类修复方式分别加以讨论。

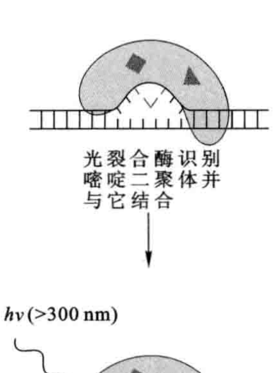
一、直接修复

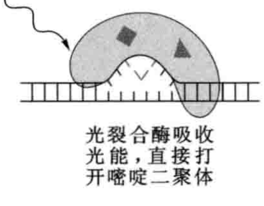
绝大多数修复系统的修复策略是将受损伤的核苷酸连同周围一些正常的核苷酸一起切除，然后，以另一条互补链上没有受到损伤的核苷酸序列作为模板，重新合成以取代原来异常的核苷酸。但直接修复与之不同，它是直接将损伤加以逆转，而不需要切除任何碱基或核苷酸。能够被这种机制修复的损伤有嘧啶二聚体和 6- 烷基鸟嘌呤。此外，DNA 链的断裂若是发生在磷酸二酯键上，且裂口是 5′- 磷酸和 3′- 羟基，细胞内的连接酶也可直接将裂口缝合。这样的修复显然也属于直接修复。

(一) 嘧啶二聚体的直接修复

参与嘧啶二聚体直接修复的酶是 DNA 光裂合酶(DNA photolyase)，它能直接识别并结合位于







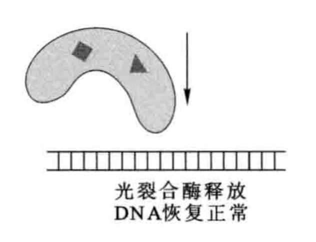


图 34-2 嘧啶二聚体的直接修复

DNA 双螺旋上的嘧啶二聚体，此过程不需要光。然后，利用它的辅基作为吸光色素捕捉光能，再用捕获的能量直接打开嘧啶二聚体，最后再与 DNA 解离（图 34-2）。

光裂合酶广泛存在于细菌、古菌和多种真核生物中，但胎盘类哺乳动物却没有这类酶。光裂合酶分为 2 类：第一类含有 2 种辅基，一种是 $FADH^-$，另外一种是 5,10- 甲炔基 - 四氢叶酸（5,10-methenyltetrahydrofolate，MTHF）或 8- 羟基 -5- 去氮黄素（8-hydroxy-5-deazaflavin，8-HDF）；另一类只含有 $FADH^-$。辅基的功能与光合作用光系统中的辅助色素相似，用来捕捉蓝光和近紫外光（350~450 nm）。

图 34-3 显示了第一类光裂合酶的辅基在直接修复嘧啶二聚体中所起的作用：首先是 MTHF 作为聚光色素，吸收波长 350~450 nm 的光能，然后将能量传递给 $FADH^-$。$FADH^-$ 正如光系统中的中心色素，接受光能以后被激发，进而失去一个电子并传递给嘧啶二聚体。嘧啶二聚体得到一个电子以后变得不稳定，随后发生一系列电子重排，直至环丁烷环被打破，电子重新回到 $FADH^·$，使 $FADH^-$ 得以再生。

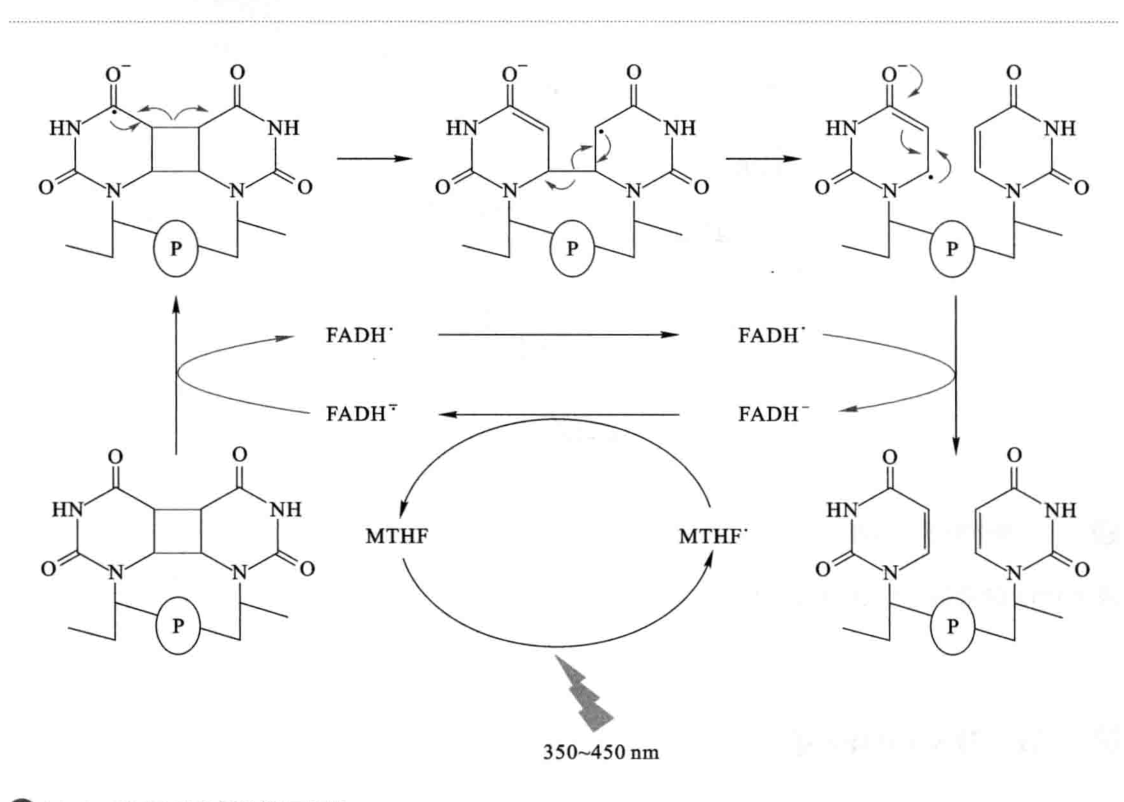


图 34-3 光裂合酶作用的详细机制

大肠杆菌的 DNA 光裂合酶既能结合单链 DNA，又能结合双链 DNA。其分子的表面在结合 DNA 的区域带有正电荷，且中间有一小孔。孔的体积和极性恰到好处，正好可以容纳嘧啶二聚体。但位于双螺旋内部的嘧啶二聚体要想结合到孔中，并与 $FADH^-$ 上的异咯嗪环接触，需要发生翻转才行。而形成的嘧啶二聚体对双螺旋结构造成的局部扭曲，降低了其与互补碱基之间的亲和力，从而使翻转变得容易。

(二) 烷基化碱基的直接修复

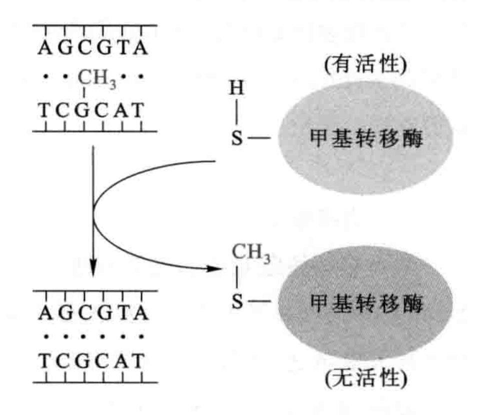


图 34-4 烷基化碱基的直接修复

催化烷基化碱基直接修复的酶是烷基转移酶（alkyltransferase），以 6- 甲基鸟嘌呤甲基转移酶（O^6-methylguanine methyltransferase，MGMT）最为常见。该酶以“自杀”的方式参与反应（图 34-4），场面颇为壮观：活性中心的 1 个 Cys 残基充当甲基受体，因此酶直接将甲基从碱基转移给这个 Cys 的巯基。一旦它得到甲基，酶也就失活了。在真核细胞中，这

框 34-1　身边的生物化学——UV 与健康

大家都知道,阳光中的 UV 对于人的健康来说是一把“双刃剑”:好的一面是它可以促进皮下的胆固醇转变为维生素 D,有利于防止维生素 D 的缺乏;不好的一面就是它对 DNA 造成的损伤。DNA 分子上相邻的嘧啶碱基,尤其是 TT,在 UV 照射下形成嘧啶二聚体或 6,4- 光产物。尽管正常的人体内有核苷酸切除修复系统,可以及时将嘧啶二聚体和 6,4- 光产物这样的损伤修复,但如果长时间受阳光的照射,UV 引发的 DNA 损伤会超过机体的修复能力,从而导致突变的发生,进而还可以诱发皮肤癌。根据美国西雅图华盛顿大学的一个研究小组 2011 年 4 月发表在美国科学院皮肤科学杂志上的一项研究报告,经常开车的美国人身体左侧得皮肤癌的机会比右侧高,而在行车方向相反的国家(如澳大利亚和英国),却是身体右侧得皮肤癌的机会大。这样的结果是不奇怪的,因为身体的一侧接受了更多的 UV 辐射所以更容易得皮肤癌。

种失活的烷基化 MGMT 随后被泛素介导的蛋白酶体降解。以牺牲一个酶分子作为代价去修复一个受到损伤的碱基,在能量学上似乎非常不经济,但在动力学上却是有利的,因为整个修复反应只有一步,可谓“一蹴而就”。这也再次验证了细胞内“遗传物质的稳定性压倒一切”。在研究此酶的基因表达的时候,发现它是一个诱导酶,显然这样的表达方式可以让“壮烈牺牲”的甲基转移酶能及时得到补充。

二、切除修复

切除修复顾名思义就是先切除损伤的碱基或核苷酸,然后重新合成正常的核苷酸,最后,再经连接酶将切口缝合。为了确保错误的碱基或核苷酸能被切掉,修复系统首先需要正确地识别受损伤的碱基或核苷酸,因此,整个切除修复过程包括识别、切除、重新合成和重新连接四大步。

切除修复又分为碱基切除修复和核苷酸切除修复,两者的主要差别在于识别损伤的机制上。前者是直接识别具体受损伤的碱基,后者并不识别具体的损伤,而主要是识别损伤对 DNA 双螺旋结构造成的扭曲。

(一) 碱基切除修复(base excision repair,BER)

BER 最初的切点是 *N*- 糖苷键,首先被切除的是受损伤的碱基。这种修复机制特别适合修复发生在碱基上对 DNA 双螺旋结构影响不大的损伤,比如尿嘧啶、次黄嘌呤、烷基化碱基、被氧化的碱基和其他一些被修饰的碱基。负责切除的酶是 DNA 糖苷酶。DNA 糖苷酶具有高度的特异性,目前已发现十多种特异性不同的 DNA 糖苷酶,但所有的 DNA 糖苷酶都是沿着 DNA 双螺旋的小沟扫描 DNA,当发现受损伤的碱基以后,立即与 DNA 结合,并导致损伤的碱基发生翻转而被挤出双螺旋,在落入酶的活性中心后被切除。碱基翻转以后留下的空间则被酶分子上一个特定的氨基酸残基(如 Arg)的侧链基团占据。

碱基切除后留在 DNA 分子上的无碱基位点(abasic site,AP 位点)并没有离开 DNA 糖苷酶的活性中心,而是很快被转移给细胞内的 AP 内切酶(AP endonuclease),以防止 AP 位点对细胞产生毒性,因为如果 AP 位点暴露在外,那无碱基的脱氧核糖环可能发生开环,而一旦开环,其 1 号位的醛基就有可能与细胞内的其他分子形成共价交联。

以尿嘧啶 -DNA 为例(图 34-5):当 U 被尿嘧啶 -DNA 糖苷酶切下以后,留下 1 个 AP 位点。AP 位点被 AP 内切酶在其上游切开,在 DNA 聚合酶的催化下,切口的 3′- 羟基端开始进行 DNA 的修补合成,合成的模板是另一条链上无损伤的互补序列。修补合成有短修补(short-patch)和长修补(long-patch)两种途径,其中短修补是主要途径,占 80%~90%,只需合成属于 AP 位点的 1 个正常的核苷酸(图中是 C)。随后,在外切酶或裂合酶或磷酸二酯酶的作用下,AP 位点被切除。最后,在连接酶的催化下,切口重新进行连接。而长修补是次要途径,占 10%~20%,作为短修补途径的备用途径,要合成 1 小段

寡聚核苷酸(2~10 nt 长),以取代原来连同 AP 位点在内的一段寡聚核苷酸。随后,外切酶将被取代的寡聚核苷酸水解。最后,同样在连接酶的催化下重新进行连接反应。细胞是选用哪一条修补途径,可能与细胞内不同类型的 DNA 聚合酶的相对活性有关。例如,真核细胞进行短修补合成只需用进行性很低的 DNA 聚合酶 β,而进行长修补合成要动用进行性较高的 DNA 聚合酶 δ 或 ε,因此如果聚合酶 β 活性很高,进行短修补的可能性就更大。

Quiz1 一种诱发癌细胞死亡的方法是让其生活在缺乏胸苷酸的状态下,试解释其中的原理。

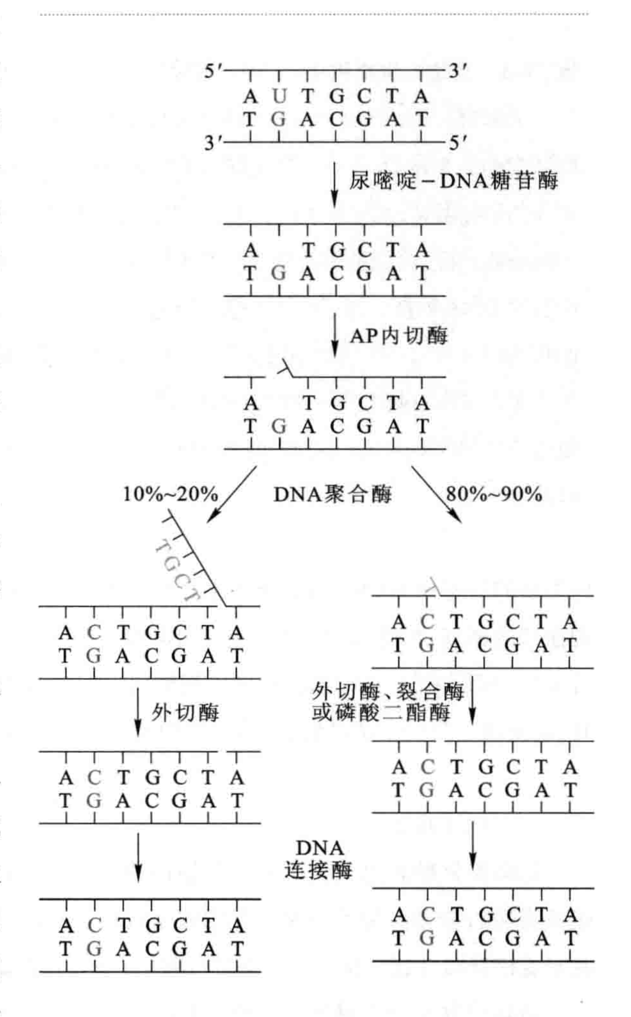


图 34-5 尿嘧啶的切除修复

除了 DNA 糖苷酶的作用可产生 AP 位点以外,自发的碱基脱落也能产生 AP 位点。后一种方式产生的 AP 位点直接由 AP 内切酶启动修复过程,其他反应与 BER 途径完全相同。

(二) 核苷酸切除修复(nucleotide excision repair,NER)

Quiz2 如何设计一个实验,确定一种生物是使用直接修复,还是核苷酸切除修复机制来修复嘧啶二聚体?

NER 最初的切点是在损伤部位附近的 3′,5′-磷酸二酯键,主要用来修复因 UV、丝裂霉素 C 和顺铂等因素造成的比较大的损伤,如嘧啶二聚体、6-4 光产物或体积较大的碱基加合物,以及链间交联导致 DNA 结构发生扭曲并影响到 DNA 复制的损伤。此外,约 20% 碱基的氧化性损伤也由它修复。在修复过程中,损伤以寡聚核苷酸的形式被切除。由于 NER 识别损伤的机制并非针对损伤本身,而是损伤对 DNA 双螺旋结构造成的扭曲,因此,NER 能够使用相同的机制和几乎同一套修复蛋白,去修复一系列性质并不相同的损伤。整个修复过程主要由 4 步反应组成:①识别损伤,由特殊的蛋白质完成,并由此引发一系列的蛋白质与受损伤 DNA 的有序结合;②切除损伤,特殊的内切酶在损伤部位的两侧切开 DNA 链,随后两个切口之间带有损伤的 DNA 片段被去除;③修复合成,DNA 聚合酶以另外一条链为模板,合成新的序列以替换已被切除的序列;④缝合切口,由 DNA 连接酶催化。

NER 又分为全局性基因组 NER(global genome NER,GGR)和转录偶联性 NER(transcription-coupled NER,TCR),前者负责修复整个基因组的损伤,速率慢,效率低;后者专门修复那些正在转录的基因模板链上的损伤,速率快,效率高。两类 NER 的主要差别在于识别损伤的机制上,至于损伤识别以后发生的修复反应并无本质上的不同。TCR 由 RNA 聚合酶识别损伤,当 RNA 聚合酶转录到受损伤部位而前进受阻的时候,TCR 系统即被启动。TCR 系统的发现使人们相信,基因模板链上出现的损伤比起基因组其他部分发生的损伤更容易得到修复。这一点是合乎逻辑的,因为模板链包含着基因组被表达的信息,它的稳定性理所当然优先受到保护。

下面就分别介绍一下细菌和真核生物的全局性 NER 系统和 TCR 系统,并就 NER 如何修复 DNA 链间交联的问题做简单的说明。

1. 细菌的 NER 系统

这里以大肠杆菌为例,它的 GGR 系统需要 UvrA、UvrB、UvrC、UvrD、DNA 聚合酶 Ⅰ/Ⅱ 和 DNA 连接酶。它们作用的具体步骤是(图 34-6):受到 ATP 水解的驱动,2 个 UvrA 与 1 个 UvrB 形成三聚体复合物($UvrA_2UvrB_1$)。该复合物与 DNA 随机结合后,沿着 DNA 链移动,以探测损伤的位置。识别损

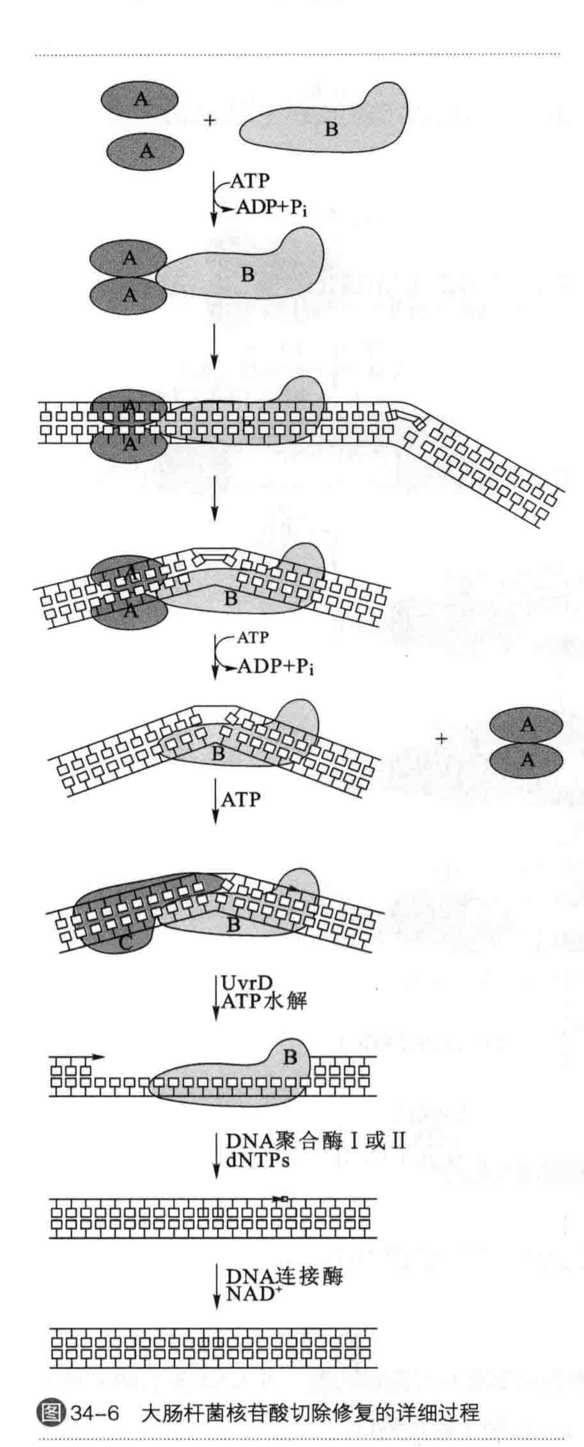


图34-6 大肠杆菌核苷酸切除修复的详细过程

伤的过程比较缓慢，为GGR系统的限速步骤。当遇到嘧啶二聚体时，通过水解ATP，造成损伤部位的DNA双螺旋发生局部解链和进一步弯曲（可达130°），致使UvrB与损伤部位发生更紧密的接触；很快，UvrA在ATP水解后离开复合物，留下UvrB横跨损伤部位；随后，UvrC被UvrB招募到损伤部位，并相互激活对方的核酸内切酶活性。被激活的UvrB在距离嘧啶二聚体下游4 nt的位置切开DNA链。与此同时，被激活的UvrC在距离嘧啶二聚体上游7 nt的位置切开DNA链。于是，一个长达13 nt的带有嘧啶二聚体的寡聚核苷酸片段产生了。随后，在解链酶UvrD的作用下，ATP被水解，包含嘧啶二聚体的DNA片段发生解链而离开双螺旋，UvrC也随之而去。最后，由DNA聚合酶Ⅰ或Ⅱ来填补空缺，DNA连接酶则缝合切口。

在TCR系统之中，RNA聚合酶充当损伤识别蛋白。一旦它进入受损伤部位就发生暂停，并将转录修复偶联因子（transcription repair coupled factor，TRCF）招募上去。TRCF的作用是促进RNA聚合酶和还没有完成的转录产物释放出来，同时将$UvrA_2UvrB_1$招募到受损伤部位，还能帮助UvrA与UvrB解离，以加快UvrB-DNA预剪切复合物的形成。余下的反应与GGR如出一辙。

2. 真核生物的NER系统

与细菌的NER系统相比，真核生物的NER系统更复杂，其修复需要更多蛋白质的参与，但作用的基本原理和过程与细菌极为相似。由于许多修复蛋白是在研究着色性干皮病（xeroderma pigmentosum，XP）、柯凯因征（Cockayne syndrome，CS）和人类的毛发二硫键营养不良症（trichothiodystrophy，TTD）时发现的，因此习惯用这些疾病名称的缩写来命名相关的修复蛋白。

以哺乳动物细胞为例，其GGR系统修复的基本步骤是（图34-7）：① XPC和hHR23B组成的二聚体识别和结合损伤的DNA。② XPC/hHR23B与损伤部位的结合，进一步加重了双螺旋结构的扭曲。③ DNA双螺旋结构的进一步扭曲使得更多的修复蛋白得以“加盟”，这些蛋白质包括TFⅡH、RPA和XPA。THⅡH有两个亚基（XPB和XPD）为解链酶，它们与DNA的损伤链结合，以水解ATP为动力，协作催化损伤部位的DNA 20~30 bp的区域朝两个相反的方向解链。RPA作为真核细胞的SSB，负责与已解开的单链区域结合。XPA虽然不是解链酶，但却是解链所必需的。由于XPA能够单独地与损伤DNA优先结合，同时又能与TFⅡH和RPA相互作用，因此它很可能与XPC/hHR23B一起将TFⅡH和RPA招募到损伤部位，还可能帮助其他蛋白质正确地定位到损伤部位。④ XPG和XPF/ERCC1被招募到已解链的损伤部位，作为DNA结构特异性的内切酶，对损伤进行切除。切点位于DNA的双链区和单链区的结合部，其中XPG先切，切点在损伤部位的3′端，与损伤位点相隔2~8 nt，ERCC1/XPF后切，切点位于损伤部位的5′端，与损伤位点相隔15~24 nt。2个切点之间DNA链的平均长度为

Quiz3 XPD突变可导致三种疾病——XP、CS和TTD的发生，为什么？

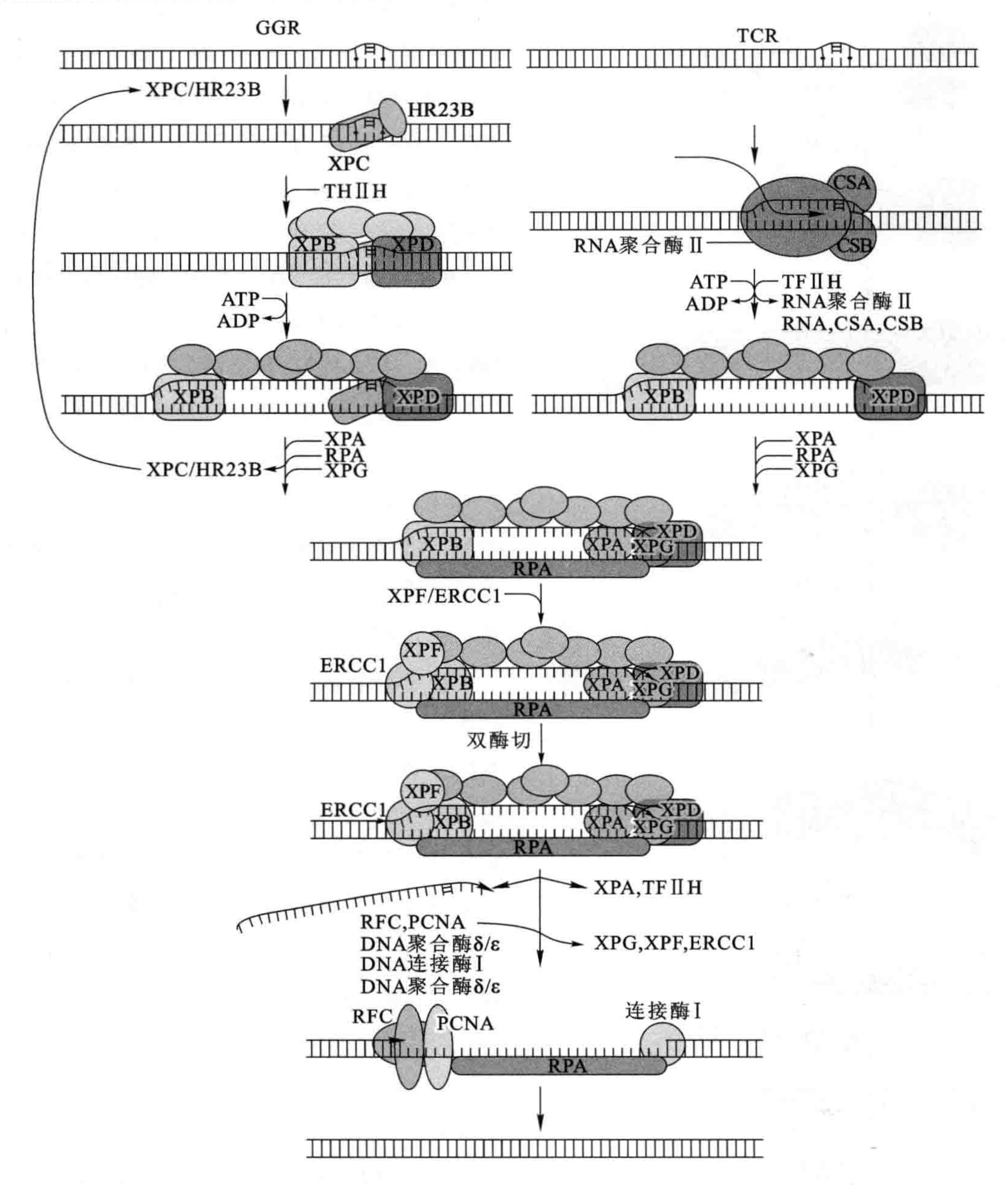


图 34-7 哺乳动物细胞的全局性 NER 和转录偶联性 NER

Quiz4 为什么 TCR 不需要 XPC？

27 nt。⑤ XPB/XPD 解链酶协助 2 个切点之间包含损伤的寡聚核苷酸的去除。⑥ DNA 聚合酶 δ 或 ε 与 PCNA 一起进行修补合成，填补空隙。⑦最后，DNA 连接酶Ⅰ缝合裂口。

如果是 TCR，就需要 XPC 以外的所有参与 GGR 的蛋白质。哺乳动物 TCR 系统的前期反应是：先是 RNA 聚合酶暂停在损伤部位，导致一小部分区域发生解链；随后，CSA 和 CSB 被招募到 RNA 聚合酶上，再帮助招募 TFⅡH、XPA、RPA 和 XPG 到损伤部位；然后，RNA 聚合酶、RNA 转录物、CSA 和 CSB 解离下来，形成与 GGR 一样的复合物，接下来的反应也就无须赘述了。

（三）错配修复（mismatch repair，MMR）

错配就是 DNA 双螺旋上的非 Watson-Crick 碱基对，产生错配的原因有 4 种：一是复制中产生的；二是 5- 甲基胞嘧啶脱氨基反应；三是不完全同源的序列发生了重组；四是 DNA 聚合酶在延伸 DNA 链时发生滑移，造成 DNA 双螺旋上出现环或突起的结构。

Quiz5 你认为还有什么方法可修复 DNA 链上的核糖核苷酸?

MMR 系统主要用来修复 DNA 分子上错配的碱基对，以及一些因“复制滑移”而诱发的核苷酸插入或缺失损伤，还能修复 DNA 复制过程中参入到 DNA 链上的少量核糖核苷酸。这里主要介绍复制产生的错配碱基是如何被修复的。

无论是哪一种生物在进行错配修复的时候，首先面临一个如何区分母链和子链的问题，因为错误

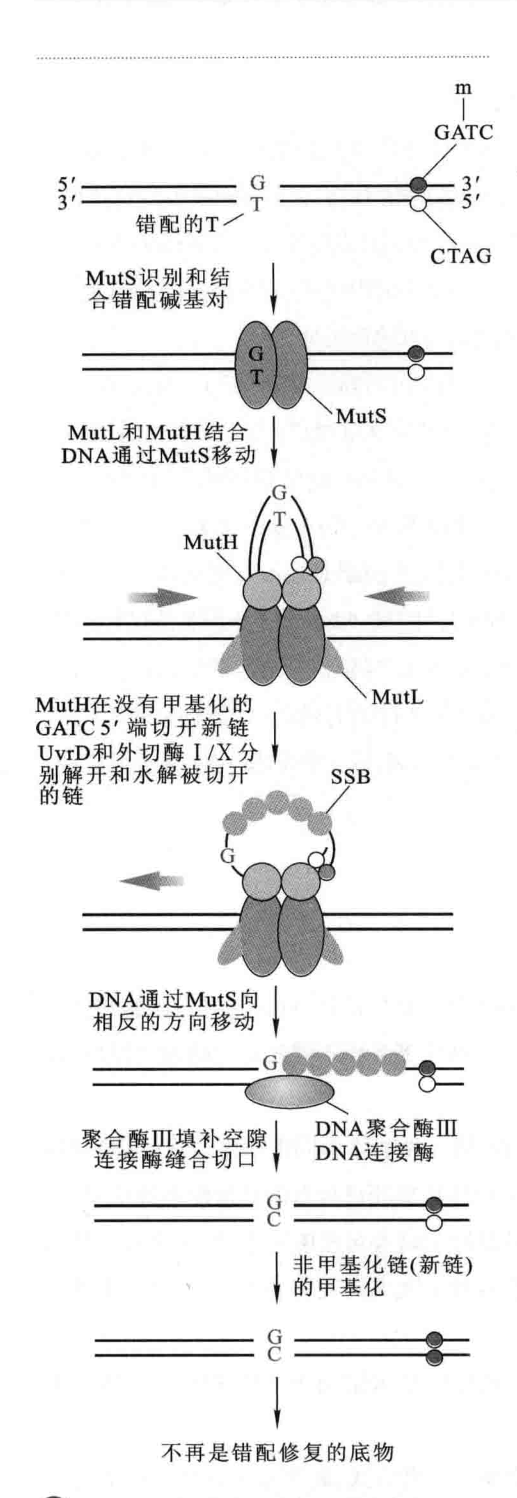


图 34-8　大肠杆菌错配修复的详细过程

只能是在子链上。错配修复系统必须能够保证只会将子链上错误的碱基切除，而不会把母链上本来正确的碱基切除。那么，MMR 系统是如何区分子代 DNA 分子中的子链和母链呢？就大肠杆菌中的这个问题，Meselson 曾大胆地作过预测，如果他是参与错配修复的酶，就会意识到刚刚复制完成的子代 DNA 分子的母链和子链的甲基化程度是不一样的，母链是高度甲基化的，子链几乎没有甲基化。后来的实验证明，大肠杆菌的 MMR 系统的确就是利用甲基化来区分子链和母链的，因此，大肠杆菌内的错配修复系统又被称为甲基化导向的错配修复(methyl-directed mismatch repair)。至于其他细菌和真核细胞是如何区分子链和母链的尚不清楚。

下面就以大肠杆菌为例，说明 MMR 系统是如何精确地修复错配碱基的。在 DNA 复制的调节机制中曾提到，新合成的子代 DNA 母链上的 GATC 序列中的 A 在 6 号位是被甲基化的，而子链在 GATC 序列上的 A 还没有甲基化，但经过一段时间以后，子链上的 GATC 在 Dam 的催化下也会被甲基化(参看第三十三章“DNA 复制”)，MMR 系统就是利用这一段时间差来工作的。参与大肠杆菌 MMR 的蛋白质除了 MutS、MutL、MutH 和 UvrD 以外，还有特殊的核酸外切酶、DNA 聚合酶Ⅲ和 DNA 连接酶。

Quiz6　如果突变的 Dam 活性大大提高，或者丧失活性，这两种情况对 DNA 复制的忠实性有何影响？

大肠杆菌 MMR 作用的主要步骤是(图 34-8)：①首先由 MutS 识别并结合除了 C—C 以外的错配碱基对或因碱基插入或缺失在 DNA 上形成的小环，MutL 随后结合；②在错配碱基对两侧的 DNA 通过 MutS 做相向移动；③ MutH、MutL 和 GATC 位点结合；④ MutH 的核酸内切酶活性被 MutS/MutL 激活，在非甲基化 GATC 的 5′端切开子链；⑤ UvrD 作为解链酶，催化被切开的含有错配碱基的子链与母链的分离，SSB 则与母链上处于单链状态的区域结合，一种特殊的外切酶将游离出来的含有错配碱基的单链 DNA 水解，如果 MutH 的切点在错配碱基的 3′端，就由外切核酸酶Ⅰ或Ⅹ从 3′→ 5′方向水解，如果 MutH 的切点在错配碱基的 5′端，就由外切核酸酶Ⅶ和 RecJ 来降解；⑥最后，DNA 聚合酶Ⅲ和连接酶分别进行缺口的修复合成和切口的缝合。

Quiz7　为什么大肠杆菌错配修复的 DNA 聚合酶是选择Ⅲ而不是其他几种？

GATC 位点与错配碱基对之间的距离可远可近，远的可达 1 kb。显然，错配的碱基离 GATC 序列越远，被切除的核苷酸就越多，重新合成子链所需要消耗的 dNTP 就越多，因此，错配修复是一个低效率和高能耗的过程。但是，无论消耗多少 dNTP，目的都是为了修复一个错配的碱基，这再次说明了在细胞内“遗传物质的稳定性压倒一切”的道理。

真核生物的 MMR 系统与细菌很接近，在真核细胞中，人们已找到与 MutS 和 MutL 同源的蛋白质，只是没有发现 MutH 的对应物。现在，有一种观点认为，子链在合成中形成的自然缺口(冈崎片段还没有连上)就是真核生物 MMR 系统用来识别子链的天然标志。

Quiz8　为什么真核细胞的错配修复没有 MutH 的同源蛋白的参与？

框 34-2　生化医药——切除修复缺陷引发的疾病以及治疗方法

着色性干皮病(XP)是由于 NER 系统有缺陷而引起的一种常染色体隐性遗传病。其主要症状是患者对紫外辐射特别敏感,患者的皮肤细胞无法修复紫外辐射造成的 DNA 损伤,会积累越来越多的突变,初起在暴露部位,如面、唇、结膜、颈部及小腿处出现雀斑和皮肤发干,类似日光性皮炎,之后出现持久性网状毛细血管扩张,暴露处皮肤色素异常,伴随角化或萎缩及癌变。XP 患者得皮肤癌的概率是正常人的千倍以上,患癌的平均年龄是 8 岁,约 2/3 的患者在进入成年以前会因皮肤癌而去世。

对于 XP,目前还没有什么特别好的治疗方法,患者若是在一出生以后就注意防护日光,那几乎完全可以避免皮肤的损伤,但这很难做到。为了防止不良症状的出现,患者必须注意严格的日光防护,特别对于紫外的防护一点都不能掉以轻心。现在已有报道,局部施用外源的 DNA 切除修复酶可用来治疗 XP。例如,有一种用脂质体包被的药物,内有重组型 T4 噬菌体的内切核酸酶 V,已通过美国 FDA 和日本有关机构的批准,患者每天可将这种药物涂敷在皮肤的暴露面,这样外源的内切核酸酶 V 可通过脂质体与皮肤细胞膜的融合,而被送入胞内。这种外源的内切酶在皮肤细胞内,可切除 DNA 分子上由紫外辐射造成的嘧啶二聚体。这种疗法将 XP 患者患癌的概率降低 30%,以及癌症发生之前的皮肤损伤降低 68%。另外,还有一种类似的药物,是用从蓝细菌提纯出来的 DNA 光裂合酶代替 T4 内切核酸酶 V,同样也取得了一定的疗效。也许,在将来是可能使用基因治疗的,根据病人有缺陷的基因,将正常的基因直接导入到皮肤细胞内,并进行表达,从而恢复患者切除修复的能力。

三、双链断裂修复

DNA 断裂特别是双链断裂是一种极为严重的损伤,如果不及时加以修复,就极易导致细胞突变或死亡。幸好,细胞针对这种损伤也有相应的修复机制。但与其他修复系统不同的是,双链断裂修复没有互补链来提供修复断裂的遗传信息。

细胞内主要存在两种不同的机制来修复 DNA 双链断裂:第一种被称为同源重组修复,它是利用细胞内一些促进同源重组的蛋白质,从姐妹染色体或同源染色体那里获得合适的修复断裂的信息,因此精确性较高;第二种机制能在无同源序列的情况下,让断裂的末端重新连接起来,因此该方式被称为非同源末端连接(non-homologous end joining,NHEJ)修复,这种方式容易发生错误,但却是人类修复双链断裂的主要方式。

(1) 同源重组修复　同源重组可用来对 DNA 双链断裂进行修复,同源重组的过程和机制详见第三十五章“DNA 重组”。

(2) 非同源末端连接修复　NHEJ 修复是最简单和最常用的一种方式,缺乏这种修复方式的突变细胞对离子辐射极为敏感。哺乳动物细胞参与这种修复方式的主要成分参看表 34-1。

► 表 34-1　参与真核细胞双链断裂修复的主要蛋白质及其功能

蛋白质	功能
Ku70	协同 Ku80 一道结合 DNA 末端,招募其他蛋白质
Ku80	协同 Ku70 一道结合 DNA 末端,招募其他蛋白质
DNA-PK_{CS}	依赖于 DNA 的蛋白质激酶的催化亚基,激活 Artemis
Artemis	受 DNA-PK_{CS} 调节的核酸酶,参与 DNA 末端的加工,使得末端适于连接
XRCC4	在 DNA 末端被 Artemis 加工好以后,协同 DNA 连接酶Ⅳ一道催化断裂的双链 DNA 分子重新连接
DNA 连接酶Ⅳ	在 DNA 末端被 Artemis 加工好以后,协同 XRCC4 一道催化断裂的双链 DNA 分子重新连接

哺乳动物细胞 NHEJ 修复的基本步骤如下(图 34-9):① Ku70 与 Ku80 形成的异源二聚体与 DNA 断裂末端结合;②两个 Ku70/Ku80 二聚体之间的相互作用将因断裂而分开的两段 DNA 强拉到一

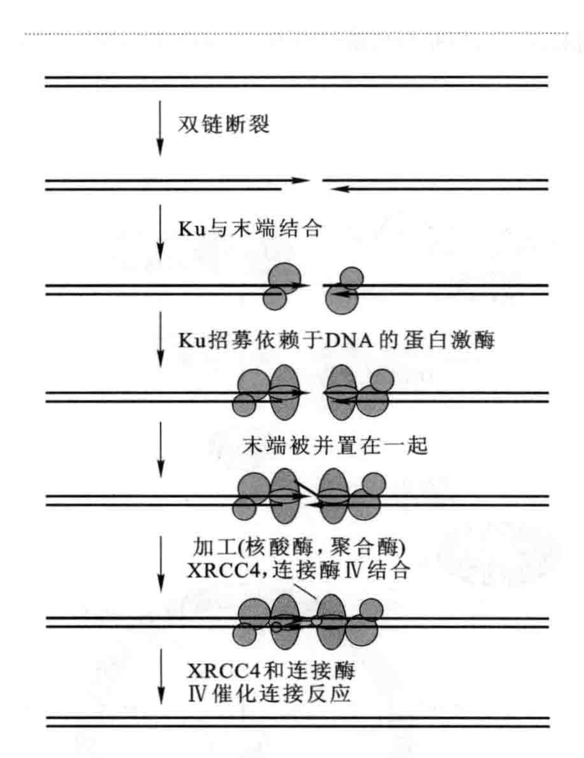


图 34-9 哺乳动物细胞 DNA 双链断裂的非同源末端连接

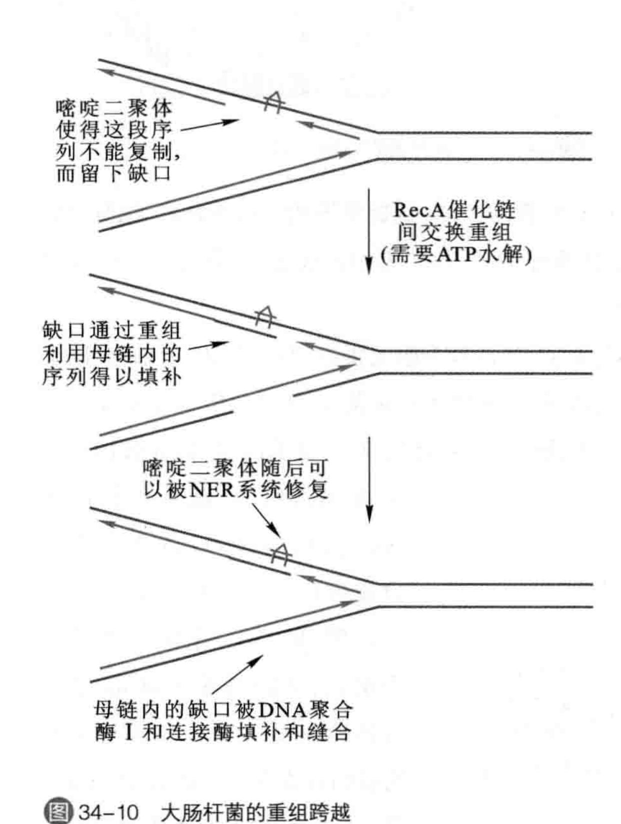


图 34-10 大肠杆菌的重组跨越

起；③ Artemis 作为 DNA-PK_{CS} 的底物与 DNA-PK_{CS} 结合，一起被 Ku70/Ku80 招募到 DNA 末端；④ DNA-PK_{CS} 一旦与 DNA 末端结合后，其蛋白质激酶的活性被激活，随后，催化 Artemis 的磷酸化修饰；⑤磷酸化后，Artemis 的核酸酶活性被激活，开始加工 DNA 的末端，水解末端突出的单链区域，为连接酶创造有效的底物；⑥ DNA 连接酶Ⅳ和 XRCC4 一道，催化已加工好的 DNA 末端之间的连接。

四、损伤跨越

所有的生物都会碰到这样的问题，即一个正在移动的复制叉遇到模板链上的损伤该怎么办？显然，最好的处理方法应该是利用修复机制将损伤迅速修复。然而在某些情形下，损伤可能无法修复，例如复制叉已经解开了母链，致使修复系统无法利用互补链作为修复的模板。而且，一个适度损伤的 DNA 进行复制也许对细胞有益，原因是它提供了通过同源重组让姐妹染色体作为修复模板的机会。于是，为了维持复制的连续性，细胞发展了两套相对独立的损伤跨越(damage bypass)“战术”，一套是通过重组，另一套是“跨越合成术”。这两套战术的原则都是先不管损伤，想方设法完成复制后再说。

严格地说，损伤跨越并不属于 DNA 修复，原因是损伤并没有完全消失，至少暂时还没有消失。然而，它的确是整个细胞面对 DNA 损伤做出的一种很重要的反应。

(一) 重组跨越

重组跨越(recombinational bypass)是利用同源重组的方法将 DNA 模板进行交换，以避免损伤对复制的抑制，从而使复制能够继续下去，而随后的复制仍然使用细胞内高保真的聚合酶，因此忠实性并无下降，故此途径被视为一种无错的系统。

以大肠杆菌为例(图 34-10)，其重组跨越的基本步骤是：①当复制叉前进到损伤位点(如嘧啶二聚体)的时候，DNA 聚合酶Ⅲ停止移动，并与模板解离，然后在损伤点的下游约 1 kb 的地方重新启动 DNA 复制，从而在子链上留下一段缺口；②在 RecA 蛋白的催化下，原来 DNA 双螺旋母链上含有的与缺口上缺失的序列一样的片段被重组到子代 DNA 上，而在母链上创造出新的缺口，但由于重组过程中的交叉是错开的，因此仍然在子链位于嘧啶二聚体的下游留下一个小的缺口；③ DNA 聚合酶Ⅰ很容易将上述缺口进行填补，而连接酶则将留下的切口缝合。

因此可见，重组跨越虽然解决了DNA不能复制的问题，但是损伤依然存在，不过留下的损伤最终可以被细胞内真正的修复系统所修复。

（二）跨越合成

跨越合成（bypass synthesis）又称为（translesion synthesis，TLS）跨损伤合成，由专门的DNA聚合酶替换停留在损伤位点上原来催化复制的DNA聚合酶，在子链上（模板链上损伤碱基的对面）插入正确或错误的核苷酸，结果导致对损伤位点无错或易错的跨越。据估计，人细胞参与TLS的DNA聚合酶至少有30多种。在结构上，这些参与TLS的DNA聚合酶与传统的DNA聚合酶有很大的差别，进行性都很低（一个或几个核苷酸），一般无校对能力。有研究表明，人体若缺失参与嘧啶二聚体无错跨损伤合成的DNA聚合酶η，很容易生癌。

1. 以大肠杆菌为代表的原核细胞的跨越合成

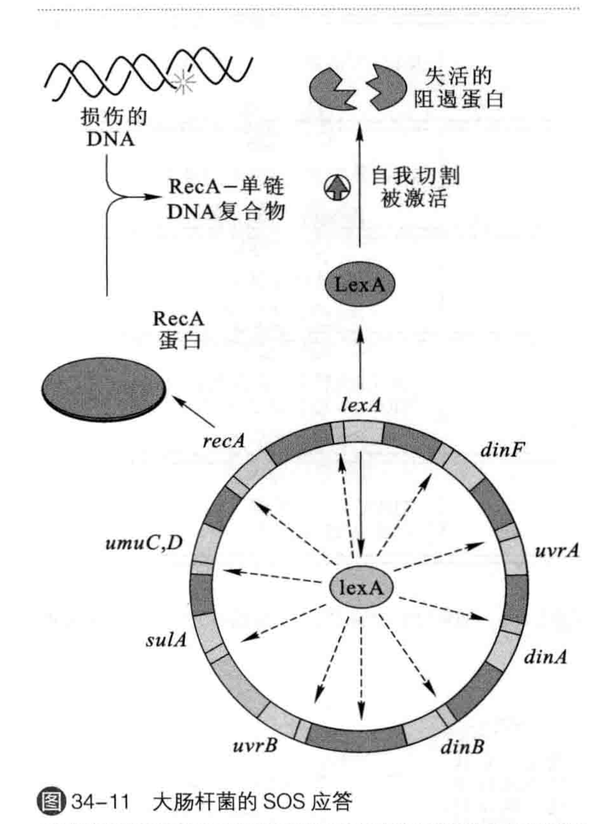


图34-11　大肠杆菌的SOS应答

大肠杆菌的跨损伤合成是其SOS应答的一部分，因此是一个可诱导的过程。SOS应答是指细胞在受到潜在致死性压力下，例如UV辐射、胸腺嘧啶饥饿、DNA修饰物的作用和DNA复制必需基因失活的情况下，做出的有利于细胞生存的代谢预警反应，包括易错的跨损伤合成、细胞丝状化（细胞伸长，但不分裂）和切除修复的激活，其中涉及近43个“*sos*”基因的表达，整个应答受到LexA阻遏蛋白和RecA激活蛋白的双重调节（图34-11）。

当大肠杆菌处于正常的生存条件下，其内部的LexA蛋白作为阻遏蛋白，与位于43个*sos*基因上游的操纵基因（一致序列为CTG-N_{10}-CAG）结合，从而阻止这些基因的表达，其中包括*lexA*和*recA*基因；当细胞面临致死性压力，其DNA遭遇到严重的损伤而出现单链缺口时，细胞内的RecA蛋白会与单链DNA结合，随后再作用于LexA蛋白，致使LexA蛋白发生自我剪切。一旦LexA蛋白发生自切割，即与*sos*基因的操纵基因解离，从而解除了对*sos*基因表达的抑制。43个*sos*基因中与跨损伤合成有关的是*dinB*、*umuC*和*umuD*，它们表达的产物分别是DNA聚合酶Ⅳ、UmuC和UmuD，UmuD受到LexA的切割变成UmuD′。当1分子UmuC与2分子UmuD′结合在一起，即组装成DNA聚合酶Ⅴ（图34-12）。

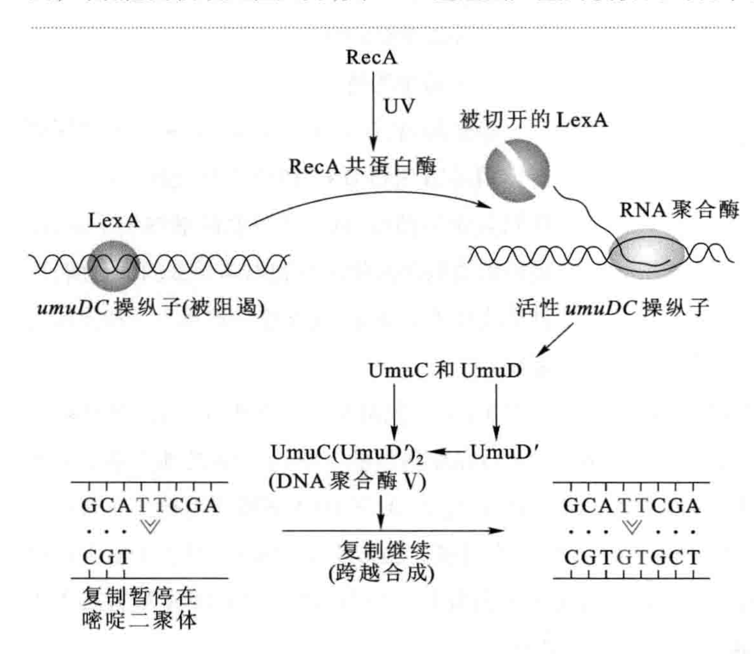


图34-12　大肠杆菌DNA损伤的跨越合成

DNA聚合酶Ⅴ在催化的时候有点像闭着眼睛摸彩，在损伤部

位因缺乏可靠的模板指导，随机地将 dNTP 参入到 DNA 链上，以克服损伤对 DNA 复制的阻碍。然而，这种盲目性是以牺牲忠实性作为代价的。参入的核苷酸错配的可能性很大，但却为细胞赢得了生存下来的机会。能被聚合酶Ⅴ跨越过的损伤包括嘧啶二聚体和 AP 位点。既然由聚合酶Ⅴ催化的 DNA 跨越合成是易错的，它就为生存下来的细胞带来了各种突变，并成为 DNA 损伤试剂诱导大肠杆菌发生突变的主要原因。

Quiz9 若是让 UmuC 呈组成型表达，则对大肠杆菌的突变有何影响?

2. 真核细胞的跨越合成

真核细胞跨越合成的方式有无错途径和易错途径两种，细胞究竟选用哪一种方式一方面取决于损伤的类型，另一方面取决于细胞内参与 DNA 跨越合成的各种聚合酶的可得性。

在易错途径中，DNA 聚合酶 ξ 和 Rev1 蛋白代替停留在嘧啶二聚体上的聚合酶 δ 或 ε，进行跨损伤合成。酵母的聚合酶 ξ 可以胜任各种损伤的跨越合成。在体内，Rev1 在损伤处插入第一个核苷酸从而启动跨损伤合成，随后，再由聚合酶 ξ 合成几个核苷酸，最后由聚合酶 δ 或 ε 取代聚合酶 ξ 和 Rev1 蛋白，继续进行 DNA 复制。在无错途径中，由 DNA 聚合酶 η 代替 DNA 聚合酶 δ 或 ε 进行跨损伤合成，在嘧啶二聚体的对面插入两个正确的 A。

体外实验表明，DNA 聚合酶 ξ 和 η 与 DNA 底物的亲和力有限，在插入一个核苷酸以后即与 DNA 解离，这种性质使得正常的聚合酶和辅助蛋白在跨损伤合成完成以后，很容易取代它们继续进行 DNA 复制。

第三节　DNA 的突变

正如本章前两节所述，DNA 会遭遇到各种各样的损伤，虽然细胞内存有多种不同的修复系统，然而，修复系统并不是完美无缺的。修复系统的不完善为 DNA 的突变打开了方便之门，因为一种损伤如果在下一轮 DNA 复制之前还没有被修复的话，就有可能直接被固定下来传给子代，有的则通过易错的跨损伤合成，产生新的错误并最终也被固定下来。这些发生在 DNA 分子上可遗传的结构变化通称为突变（mutation）。

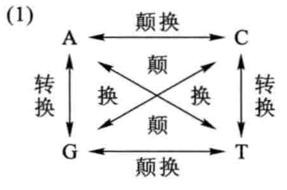
一、突变的类型与后果

DNA 突变的本质是其核苷酸序列发生的变化。根据核苷酸序列的变化方式，DNA 突变可分为点突变（point mutation）和移框突变（frameshift mutation）。

（一）点突变

点突变又被称为碱基对置换（base-pair substitution），它是指 DNA 分子某一位点上所发生的一种碱基对变成另外一种碱基对的突变，可分为转换（transition）和颠换（transversion）两种形式（图 34-13）。其中，转换是指两种嘌呤碱基（A 和 G）或两种嘧啶碱基（T 和 C）之间的相互转变，颠换是指嘌呤碱基和嘧啶碱基之间的互变。

(1)



(2)

沉默突变

TGT → TGC

Cys → Cys

错义突变

TGT → TGG

Cys → Trp

无义突变

TGT → TGA

Cys → 终止

图 34-13　碱基突变的几种方式

点突变的后果取决于它的位置和具体的突变方式。如果发生在垃圾 DNA（junk DNA）上，就可能不会产生任何后果；如果发生在一个基因的启动子或者其他调节基因表达的区域，可能就会改变基因表达的效率；如果发生在一个基因的内部，情况就比较复杂，一方面取决于突变基因是蛋白质基因还是非蛋白质基因，另一方面如果是蛋白质基因，就取决于究竟发生在它的非编码区，还是编码区。这里只介绍发生在蛋白质基因内部的点突变。

如果突变发生在蛋白质基因的非编码区，就可影响到该蛋白质基因的转录、转录后加工和翻译等；如果发生在蛋白质基因编码区，就会有三种不同的后果：

(1) 突变的密码子编码同样的氨基酸，这样的突变对蛋白质的结构和功能不会产生任何影响，因

此被称为沉默突变（silent mutation）。

(2) 突变的密码子决定不同的氨基酸，这样的突变可能对蛋白质的功能不产生任何影响或影响微乎其微，也可能产生灾难性的后果。由于出现了错误的氨基酸，这样的突变被称为错义突变（missense mutation）。如果突变的氨基酸与原来的氨基酸具有同种性质，这种突变就称为中性突变（neutral mutation）。中性突变对蛋白质的功能影响一般很小。如果突变的氨基酸与原来的氨基酸性质差别较大，这样的错义突变通常就是灾难性的，是许多疾病发生的分子基础，如镰状细胞贫血和囊性纤维变性（cystic fibrosis）。

(3) 突变的密码子变为终止密码子或者相反。前一种情况可导致一条多肽链被截短，这被称为无义突变（nonsense mutation），而后一种情况会加长一条多肽链，被称为加长突变（elongation mutation）或通读突变（read-through mutation）。无义突变究竟会给一个蛋白质的功能带来什么影响，主要取决于丢掉了多少个氨基酸残基。

（二）移框突变

移框突变又称为移码突变，它是指在一个蛋白质基因的编码区发生的一个或多个核苷酸（非 3 的整数倍）的缺失或插入。由于遗传密码是由 3 个核苷酸构成的三联体密码，因此，这样的突变将会导致阅读框发生改变，致使突变点下游的氨基酸序列发生根本性的改变，但也可能会引入终止密码子而使多肽链被截短。移框突变究竟对蛋白质功能有何影响，取决于突变点与起始密码子的距离。显然，离起始密码子越近，功能丧失的可能性就越大。

（三）隐性突变和显性突变

Quiz10 预测以下突变哪些是隐性的，哪些是显性的？原癌基因、抗癌基因、XP 蛋白、β 珠蛋白。

对于真核生物而言，DNA 突变可能是隐性的（recessive），也可能是显性的（dominant）。如果突变仅仅是灭活一种蛋白质，那么这种突变就一般产生隐性性状，这是因为染色体是成对的（同源染色体），每一个基因至少有 2 个拷贝，一条同源染色体上正常的基因能够抵消另一条同源染色体上突变的基因对细胞功能和性状的影响，只有一对同源染色体上两个相同的基因都发生突变，才会影响到表型；如果突变产生的蛋白质对细胞有毒，这种毒性就无法被另外一条染色体上正常基因表达出来的正常蛋白质所抵消或中和，这种突变就表现为显性。显性突变只需要两条同源染色体上任意一个拷贝的基因发生突变，就可以带来突变体的表型发生变化。

二、突变的原因

几乎任何导致 DNA 损伤的因素都能导致 DNA 突变，前提是造成的损伤在 DNA 复制之前还没有被修复。因此可以这样认为，导致 DNA 损伤的原因在某种意义上就是导致 DNA 突变的原因。正如 DNA 的损伤有内外两种因素一样，DNA 的突变也是如此，由内在因素引起的突变称为自发突变（spontaneous mutation），由外在因素引发的突变称为诱发突变（induced mutation）。各种导致 DNA 突变的内外因素统称为突变原（mutagen）。

（一）自发突变

1. 自发点突变

导致自发点突变的原因有：

(1) DNA 复制过程中的错配　关于导致错配的原因在损伤一节已有介绍，这里仅仅说明碱基的互变异构对碱基转换形成的贡献。DNA 分子的碱基可以自发地进行互变异构(参看第五章“核苷酸”)，例如，T 有酮式和烯醇式，酮式碱基和烯醇式碱基因氢键受体或供体的变化，故在配对性质上会有差异，如烯醇式的 T 可以和 G 配对。试想一下，如果 DNA 复制正好遇到模板链上的 1 个烯醇式 T，G 就代替 A 与 T 配对，从而形成错配的 T : G 碱基对。倘若这个错配的碱基对逃过了聚合酶的自我校对以及错配修复系统，那么在经过又一轮复制以后，最终将导致原来的 T : A 碱基对转换为 C : G 碱基对。

Quiz11 假定DNA分子的第四个碱基是U(不是T),那么DNA分子中的CpG二核苷酸序列发生甲基化后,对DNA的突变率还有没有影响?为什么?

(2) 自发脱氨基　胞嘧啶和5-甲基胞嘧啶经脱氨基反应分别转变成U和T,由于细胞内的BER系统很容易识别和修复DNA分子上的U,因此由胞嘧啶脱氨基引发的突变的可能性极小。但是,5-甲基胞嘧啶就不一样了,在人类基因组之中,它通常出现在CG序列上。如果它脱氨就变成了T,由于T是DNA分子上正常的碱基,机体没有专门的修复系统纠正这种错误,那在经过一轮DNA复制以后,就会导致C:G碱基对转换为T:A碱基对(图34-14)。

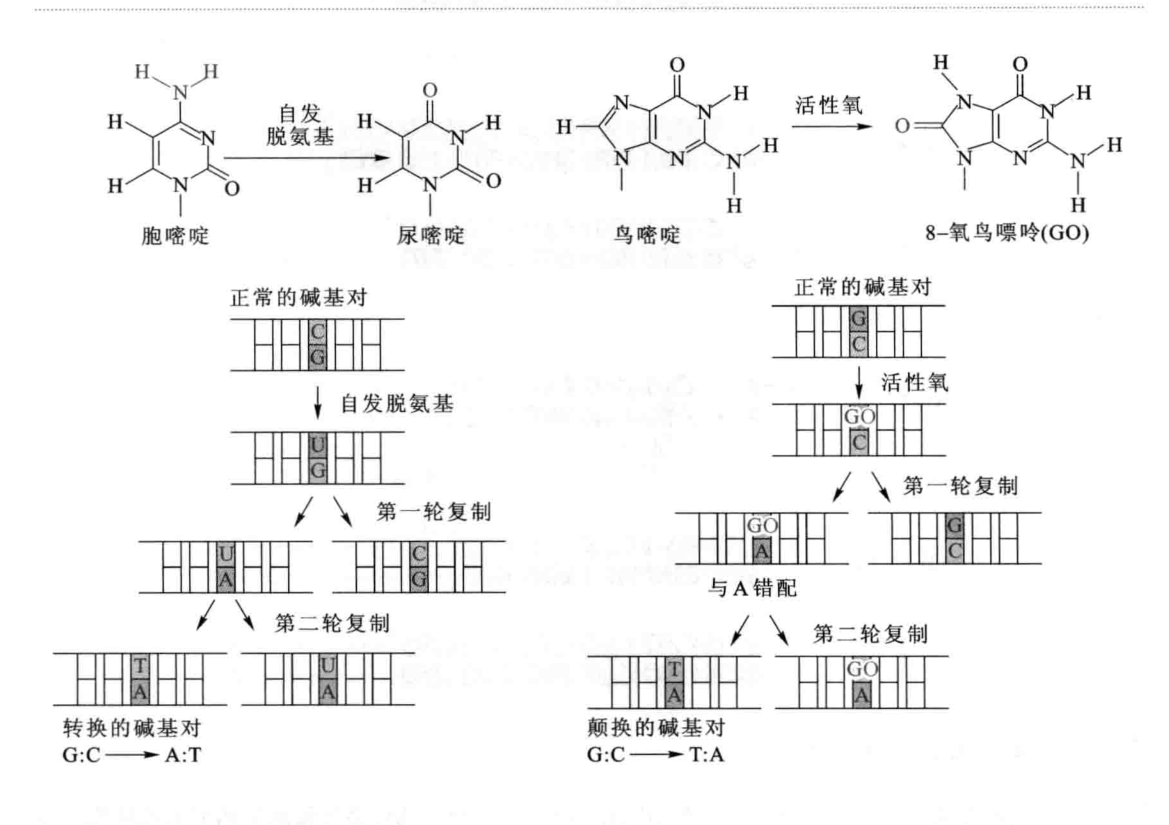


图34-14　自发脱氨基和活性氧作用引起的碱基转换(Hartwell等,2011)

(3) 活性氧的氧化　细胞的正常代谢产生的活性氧会对碱基造成损伤,这些损伤能够改变碱基的配对性质。如活性氧作用鸟嘌呤的产物8-氧鸟嘌呤与A配对,这可以导致G:C碱基对到T:A碱基对的颠换(图34-14)。

(4) 碱基的烷基化　这里是指细胞内一些天然的烷基化试剂,如SAM,引起的DNA上某些碱基的甲基化,而改变了碱基的配对性质。

2. 自发的移框突变

引起自发移框突变的主要原因有"复制滑移"(replication slippage)和转座作用。

(1) 复制滑移　"复制滑移"通常出现在一些具有短重复序列的部位。在这样的部位,子链和母链之间容易发生错配而形成突环结构。如果突环出现在子链上,复制就会向后滑移,导致插入突变;如果突环出现在母链上,复制就会向前滑移,导致缺失突变(图34-15)。如果这种突变发生在一个基因的编码区,就可能产生异常的蛋白质而导致机体病变,如亨廷顿病。

(2) 转座作用　转座子指是细胞内可移动的DNA片段。当转座因子插入一个基因内部以后,不仅会引起移框突变,还可能导致基因的中断和失活等其他变化。

(二) 诱发突变

1. 诱发点突变

能够诱发点突变的试剂有以下几类:

(1) 碱基类似物　如5-溴尿嘧啶(5-BrU)和2-氨基嘌呤。碱基类似物进入细胞后,很容易通过补救途径转变成相应的dNTP类似物,然后在DNA复制的时候,代替正常的dNTP进入DNA链。但

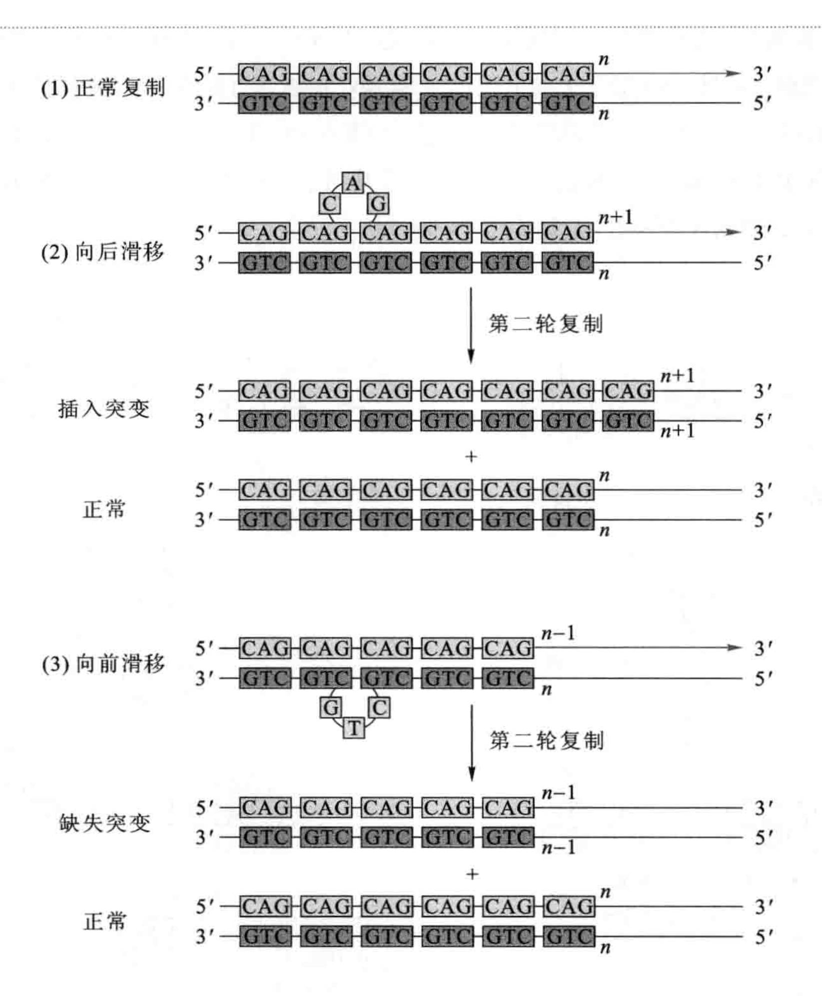


图 34-15　复制滑移引起的移框突变

是，它们毕竟在结构上与真正的碱基有差异，因此一旦参入到 DNA 链，会导致碱基的配对性质发生变化。以 T 的类似物 5-BrU 为例，当它代替 T 参入到一个正在合成的 DNA 链以后，由于在体内转变为烯醇式的机会提高，而烯醇式的 5- 溴尿嘧啶将与 G 配对，于是 5-BrU 的参入最终会导致 DNA 分子中的 A：T 碱基对转换为 G：C 碱基对。

同理，2- 氨基嘌呤作为 A 的类似物，如果参入到 DNA 分子之中，在复制中就能够与 T 或 C 配对，这样的配对最终可导致 A：T 到 G：C 或 G：C 到 A：T 的转换。

(2) 烷基化试剂　如氮芥和硫芥等。烷基化试剂能够修饰碱基而改变碱基的配对性质，从而将碱基对的转换引入 DNA 分子之中。例如，6- 甲基鸟嘌呤可以和 T 配对，致使 G：C 转换为 A：T。此外，某些双功能烷基化试剂可导致 DNA 的链间交联，而引起染色体的断裂。

(3) 脱氨基试剂　如亚硝酸和亚硫酸。亚硝酸能加快碱基的自发脱氨基作用。C、A 和 G 在亚硝酸的作用下，分别转变成 U、I 和黄嘌呤。除了黄嘌呤的配对性质与 G 一样以外，其他两种碱基配对性质都有变化，这种变化将最终导致碱基对的转换（图 34-16）。亚硫酸只作用于 C，促进 C 脱氨基转变成 U，因此可以导致 C：G 碱基对转换成 T：A 碱基对。

(4) 羟胺　羟胺在体内可直接修饰碱基而改变碱基的配对性质，从而诱发碱基对的转换，如 C 经羟胺的修饰变成能与 A 配对的羟胞嘧啶，这最终可导致 CG → TA（图 34-16）。

2. 诱发移框突变

嵌入试剂（intercalating agent），如吖啶黄（acridine orange）、原黄素（proflavin）和 EB 等，是一类扁平的多环分子。它们能够与 DNA 分子上的碱基相互作用，插入到碱基之间（图 34-17）。这样的插入将拉长 DNA 双螺旋，并骗过 DNA 聚合酶，致使 DNA 在复制的时候，发生移框突变。如果嵌入试剂插入到母链上，就会在子链上嵌入分子的对面部位随便插入一个核苷酸，造成插入突变；相反，如果嵌入

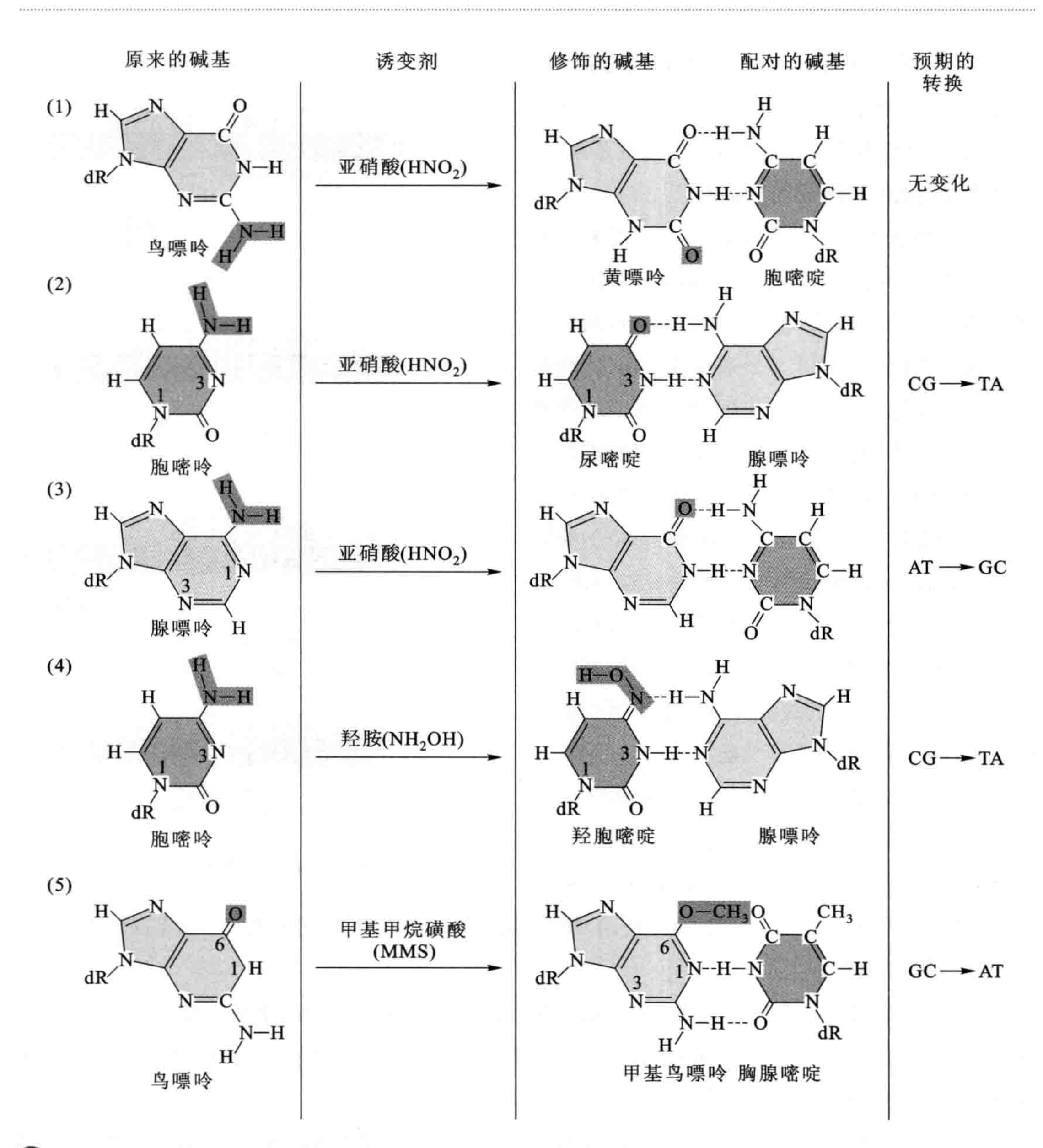


图 34-16 诱变剂诱发的点突变(Russell,2009)

分子插入到一个正常延伸的子链上,那在进行下一轮复制的时候,当嵌入分子丢失以后,就会导致缺失突变。

除了上述各种能够直接导致 DNA 分子发生突变的试剂以外,还有一些因素,特别是离子辐射和 UV,通过损伤 DNA 诱发易错的跨损伤合成和 NHEJ 而导致突变。

三、回复突变与突变的校正

(一) 回复突变

DNA 突变并不是不可逆转的,如果在老的突变位点上发生第二次突变,致使原来的表型得到恢复,这样的突变就称为回复突变(back mutation)。表型能够在回复突变中恢复的可能原因是:突变点编码的氨基酸变成原来的氨基酸或性质相似的氨基酸,从而使原来突变蛋白的功能得到全部或部分恢复。

(二) 校正突变

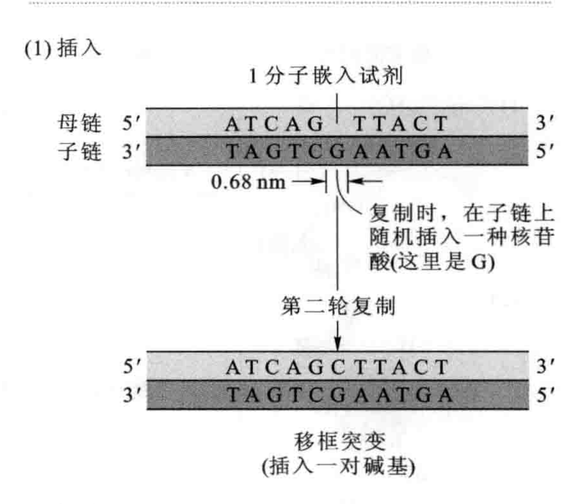
校正突变(suppressor mutation)是指发生在非起始突变位点上,但能够掩盖或抵消起始突变的第二次突变,它分为基因内校正和基因间校正。

1. 基因内校正

基因内校正与起始突变发生在相同的基因内(图 34–18),它可能是通过点突变或移框突变来实现校正,不过点突变一般只能通过点突变来校正,移框突变只能通过移框突变来校正。

如果是通过点突变来校正,一般就是通过恢复一个基因产物内 2 个残基(氨基酸残基或核苷酸残基)之间的功能联系来实现。具体机制可能是两次突变相互抵消了 2 个残基的变化,从而恢复了 2 个残基之间的相互作用,致使基因产物能够正确地折叠,或者使 2 个相同的亚基能够组装成有功能的同源二聚体。现举一例说明,假如一个蛋白质的正确折叠需要在 Lys3 和 Glu50 的侧链之间形成盐键,显然,如果 Lys3 突变成 Glu3,就会导致原来的蛋白质不能正确折叠而丧失功能,但如果 Glu50 发生第二次突变变成了 Lys50,就可以恢复 Glu 和 Lys 之间的盐键,致使突变的蛋白质仍能正确折叠,并具有原有的功能。

如果是通过移框突变来校正,那起始突变一般也是移框突变,且移框的方向相反,数目相同。例如,一个基因的第一次突变是 +1 移框,如果有第二次突变正好发生在它的附近,而且是 –1 移框的话,第二次突变很有可能就是一次基因内校正。



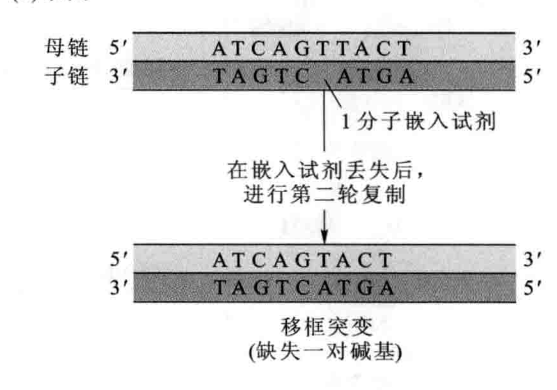


图 34–17 嵌入试剂诱发的移框突变(Russell,2009)

2. 基因间校正

基因间校正发生在与第一次突变不同的基因上,绝大多数是在翻译的水平上起作用。这种发生第二次突变具有校正功能的基因称为校正基因。每一种校正基因只能作用一种类型的突变,即是无义突变、错义突变和移框突变中的一种。

校正基因一般是通过恢复 2 个不同的基因产物之间(2 条不同的多肽链、2 个不同的 RNA 或者 1 条多肽和 1 分子 RNA)的功能关系来实现的。校正基因通常编码 tRNA,因为它们是通过反密码子与 mRNA 上的密码子相互作用来参与翻译的,所以发生在 tRNA 反密码子上的突变可用来校正 mRNA 上的一个密码子的突变,恢复密码子和反密码子之间的互补关系,从而使翻译出来的蛋白质的氨基酸序列恢复如初。

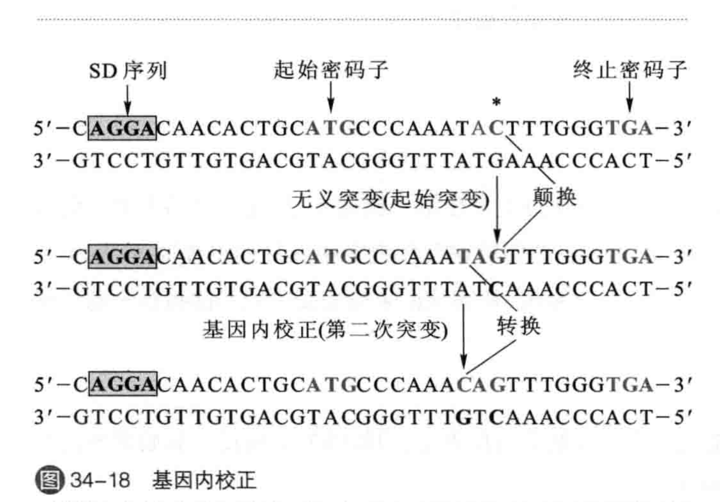


图 34–18 基因内校正

校正 tRNA 除了能够校正无义突变以外,还能校正错义突变,甚至移框突变。一种校正 tRNA 究竟能够校正何种突变完全取决于反密码子发生什么样的突变。例如,有一种校正 tRNA 的反密码子由 4 个核苷酸组成,显然,这种校正 tRNA 可以校正 +1 移框突变。

在细胞内,校正 tRNA 的基因与野生型 tRNA 的基因共存,其产物即校正 tRNA 会与野生型 tRNA 或翻译的终止释放因子竞争,这可能会导致正常的

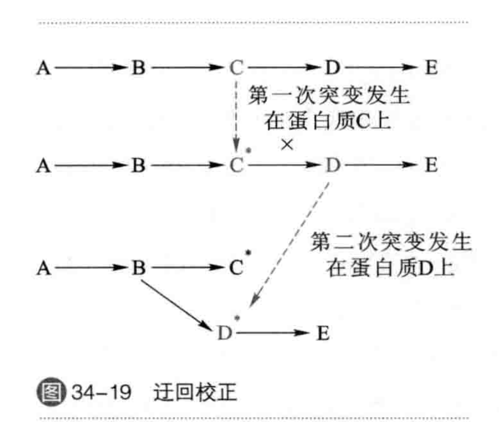


图 34-19 迂回校正

翻译发生错义或通读。

如果校正基因是一个蛋白质基因，校正机制一般就是通过它编码的蛋白质上的一个氨基酸残基的变化，抵消了发生第一次突变的那一个蛋白质上的氨基酸变化，从而使它们仍然能够正常地组装在一起形成有功能的异源寡聚体蛋白。除此以外，还有一种叫迂回校正（bypass suppressor）的机制，该机制通常适用于一条信号通路。如图 34-19 所示，蛋白质 C 的突变使得信息无法从 C 传给 D，从而导致整个信号通路无法正常运转。然而，蛋白质 D 的同时突变使得它能够绕过 C，直接从蛋白质 B 得到信息，从而使原来的信号通路恢复畅通。

四、突变原与致癌物质之间的关系以及致癌物质的检测

据估计，多达 80% 的人类癌症是由各种导致 DNA 损伤或者干扰 DNA 复制的致癌物质引发的，因此致癌物质一般也是突变原。由于许多致癌物质是人工合成的物质（如杀虫剂和农药），因此需要建立一套快速检测一种物质是否是致癌物质的方法。既然致癌物质一般是突变原，那么完全可以根据一种物质的致变性来推测其潜在的致癌性。

Bruce Ames 于 1975 年建立的沙门氏菌回复突变试验法，即 Ames 试验（Ames test）法就是使用突变性推测致癌性的一种较为流行的检测方法。该法快捷、简便、敏感、经济，且适用于测试混合物，反映多种化学物质的综合效应。

e34-1 Ames 试验法

科学故事——TTDA 的发现

某些基因的遗传性疾病是很罕见的，原因是它们影响到了细胞最重要的功能，从而导致突变体无法生存。只有有限的突变不会影响到细胞内不可缺少的功能，这些突变体能够生存下来，并供人类研究。

1968 年，James Cleaver 在 *Nature* 上发表了一篇关于 3 个罕见的着色性干皮病患者（XP）细胞内 DNA 修复缺陷的论文。据 Cleaver 回忆，他是于 1967 年 4 月，在当月的旧金山年历（San Francisco Chronicle）里，看到了一篇描述一位 XP 患者因为阳光诱发遗传性皮肤癌的文章而受到启发，猜想 XP 患者可能就是他想做的关于辐射损伤人类 DNA 研究的天然突变体。随后，他从病人体内抽取 XP 细胞，开始对 XP 与 DNA 损伤、修复的关系进行系统的研究，并取得了一系列突破性的成果。

NER 系统的功能包括切除 UV 辐射产生的嘧啶二聚体，它对于所有接触阳光的生物来说都是极为重要的。XP 患者缺乏正常的有功能的 NER 系统，因此对 UV 极度敏感，容易被阳光诱发得皮肤癌。NER 系统在各种生物（从酵母到仓鼠再到人）中都是由一套高度同源的蛋白质组成。这些蛋白质协调一致，识别 DNA 损伤，在损伤区域解开 DNA 链，切除损伤的核苷酸，填补切除后留下来的缺口。人类 XP 疾病就是参与 NER 修复的 7 种蛋白质（XPA 至 XPG）缺陷造成的。

在法国，由 Jean-Marc Egly 领导的研究小组一直在研究基础转录因子 TFⅡH。1993 年，他们报道，构成 TFⅡH 的两个成分也是 DNA 修复蛋白，即 XPB 和 XPD。这个很重要的发现将转录和 DNA 修复联系到一起，既然一种蛋白质同时在两个系统中发挥作用，就解释了为什么完全缺乏 XPB 或 XPD 的小鼠在胚胎期就死亡的现象。然而，在 *XPB* 和 *XPD* 基因上发生的某些突变也出现在一些 XP 病人身上，但病人依然能生存。后来发现，含有类似突变的小鼠也能够生存，并表现出某些 XP 患者的症状。但是，在 *XPD* 和 *XPB* 基因上有缺陷的 XP 患者的临床表型并不相同：某些病人表现出渐发性的神经退化，而其他病人则表现出身材矮小、发育迟缓和恶病质状态。表型的差异似乎反映了同一种蛋白质在不同功能上的缺陷。

不久，意大利科学家 Miria Stefanini 领导的研究小组以及荷兰和法国科学家发现，某些 XPD 或 XPB

有缺陷的病人有毛发二硫键营养不良症（TTD）的表型。这些病人的细胞在培养的时候与 XP 细胞相似。1993 年，他们发现一个 TTD 病人的细胞（TTD-A）对 UV 辐射极度敏感，但他的 XPB 和 XPD 没有任何缺陷。于是在这种 TTD-A 细胞内，共耗时 10 年的搜寻缺陷基因的"战斗"打响了！既然 XPB 和 XPD 是 TFⅡH 的组分，那么，这种复合物首先进入了科学家寻找 TTD-A 致病基因的视线。TFⅡH 的检测实验表明，TTD-A 细胞内 TFⅡH 浓度比正常细胞要低。可是，后来对每一个 TFⅡH 组分的序列分析显示一切正常，没有任何突变。而且，添加纯化的正常 TFⅡH 蛋白的每一个组分都不能纠正 TTD-A 的缺陷。就在科学家们百思不得其解的时候，酵母和藻类中的新发现引起了荷兰科学家的注意。以 Ranish 为首的研究小组在酵母中发现，TFB5（TFⅡH 的新组分）与 TTD-A 相似，缺乏 TFB5 的酵母细胞对于 UV 辐射极为敏感。此外，在莱茵衣藻（*Chlamydomonas reinhardtii*）中发现的 UV 敏感型突变体校正蛋白——Rex1 与 TFB5 存在序列的同源性！

于是，他们决定利用这个重要的信息确定人是否也含有 TFB5 的同源物，如果有，那么 TTD-A 细胞是否在 TFB5 的同源物上有缺陷？很快，一个国际研究小组成立了，小组内的专家利用细胞、分子和物理学的手段进行研究。不久，他们在人细胞内克隆出 TFB5 的类似物（被称为 TTDA），其相对分子质量为 8×10^3，并证明了它与人 TFⅡH 其他组分之间的相互作用。进一步的研究还表明，正常的 TTDA 可纠正 TTD-A 细胞的功能缺陷。那么，为什么如此重要的蛋白质经过这么久才被发现？最可能的原因是它太小了，只有 8×10^3，很容易在纯化过程中与染料一起跑出凝胶，所以被忽略了。但不管怎么样，毅力终究会带来胜利！

e34-2 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/DNA_repair（维基百科有关 DNA 修复的内容）
2. http://en.wikipedia.org/wiki/Mutation（维基百科有关突变的内容）
3. http://www.web-books.com/MoBio/Free/Ch7G.htm（免费的分子生物学在线课程，有关 DNA 损伤修复的内容）
4. http://www.wormbook.org/chapters/www_DNArepair/DNArepair.html（书虫网站提供的有关 DNA 复制的最新内容）

参考文献：

1. Knobel P A, Marti T M. Translesion DNA synthesis in the context of cancer research. *Cancer Cell International*, 11:39, 2011.
2. Shen Y, et al. Mispaired rNMPs in DNA are mutagenic and are targets of mismatch repair and RNases H. *Nature Structural & Molecular Biology*, 19:98–104, 2012.
3. Krutyakov V M. Eukaryotic error-prone DNA polymerases: The presumed roles in replication, repair, and mutagenesis. *Molecular Biology*, 40:3–11, 2006.
4. Goodman M F. Error-prone repair DNA polymerases in prokaryotes and eukaryotes. *Annu Rev Biochem*, 71: 17–50, 2002.
5. Friedberg E C, et al. Error-prone DNA polymerases: novel structures and the benefits of infidelity. *Cell*, 107:9–12, 2001.
6. Sutton M D. The SOS response: recent insights into umuDC-dependent mutagenesis and DNA damage tolerance. *Annu Rev Genet*, 34:479–497, 2000.

第三十五章　DNA 重组

任何物种的基因组 DNA 都不是固定不变或绝对稳定的。它不仅会遭受损伤，还会发生突变和重组（recombination）。DNA 重组是指发生在一个 DNA 分子内或两个 DNA 分子之间核苷酸序列的交换、重排和转移现象。它主要包括同源重组（homologous recombination）、位点特异性重组（site-specific recombination）和转座重组（transposition recombination）3 种形式。如果说突变能够为生物带来变异的话，那么重组就可以在基因组内创造不同的变异组合。生物通过重组，不仅可以产生新的基因或等位基因的组合，还可以创造出新的基因，提高种群内遗传物质的多样性；此外，重组还被用于 DNA 修复；某些病毒还利用重组将自身的 DNA 整合到宿主细胞的 DNA 上。在掌握天然重组原理的基础上，人们还可以利用重组来进行基因作图、基因敲除、基因治疗和转基因生物的培育等。

本章将重点阐述同源重组、位点特异性重组和转座重组的机制、原理和功能，至于重组的应用会在第四十四章“重组 DNA 技术”做专门介绍。

第一节　同源重组

同源重组也称为一般性重组（general recombination），它发生在含有同源序列的两个 DNA 分子之间，即在两个 DNA 分子的同源序列之间直接进行交换。进行交换的同源序列可能是完全相同的，也可能是近乎相同的。细菌的转化（transformation）、接合（conjugation）、转导（transduction）和真核细胞同源染色体之间的交换等都属于同源重组。

进行同源重组必须满足 5 个条件：①在交换区域含有相同或几乎相同的碱基序列；②双链 DNA 分子之间发生互补配对；③需要重组酶（recombinase）的催化；④形成异源双链；⑤发生联会（synapsis）。

为了解释同源重组的机理，人们在实验的基础上，提出了各种模型，下面就介绍几种具有代表性的同源重组模型。

一、同源重组的模型

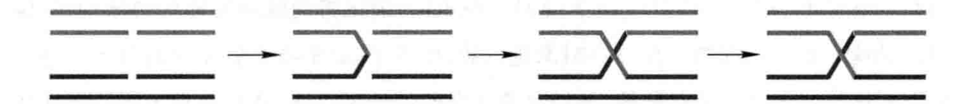
（一）Holliday 模型

此模型由 Robin Holliday 在 1964 年提出，虽然经过多次修改，但其核心内容仍然是当今各种解释同源重组模型的基石。最初的 Holliday 模型的主要内容是：

（1）2 个同源的 DNA 分子并列在一起。

（2）2 个 DNA 分子各有一条链在相同的位置被切开，被切开的链必须具有相同的极性。

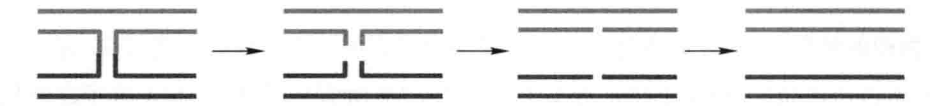
（3）被切开的链进行交换并与同源的链连接，形成 X 状的 Holliday 连接（Holliday junction）或 Holliday 结构（Holliday structure）或 Holliday 中间体（Holliday immediate）。



（4）1 个分子相对于另外 1 个分子作 180° 的旋转，得到另一种 Holliday 结构。



(5) Holliday 连接的分离。有两种分离的方法,第一种方法是相同的链被第二次切开,结果产生与原来一模一样的两个非重组的 DNA。



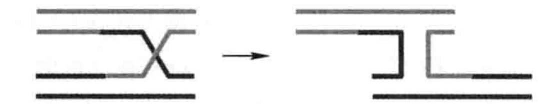
第二种分离的方法是另一条链被切开,然后经过重新连接,产生重组 DNA。



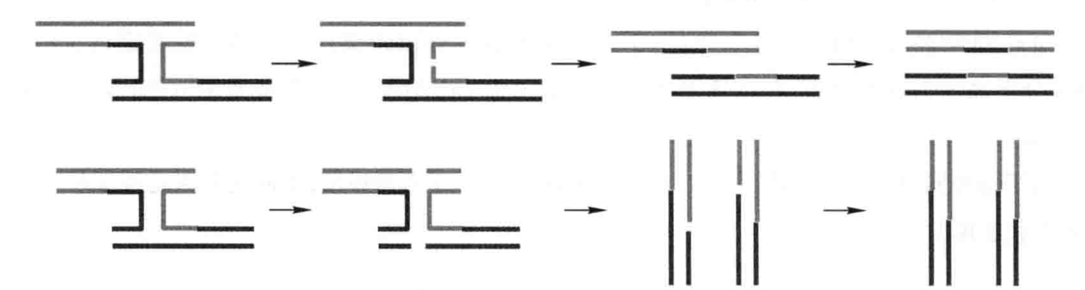
由于上述模型过于简单,无法解释许多天然的同源重组现象,人们很快就对此模型进行了修正。首先在 Holliday 连接形成之后引入 1 个新的步骤,即分叉迁移(branch migration)。如下图所示,Holliday 连接形成以后,其分叉可以向两边移动,这样的移动将会导致 1 个 DNA 分子上一条链上的部分序列转移到另 1 个 DNA 分子之中。



上述经迁移过的 Holliday 连接,再通过内部 180° 旋转可得到以下结构:



最后的分离仍有两种方法,但分叉的迁移导致了非重组的 DNA 分子也含有异源双链。



(二) Holliday 模型的实验证据

1976 年,David Dressler 和 Hunt Potter 发表了一系列支持 Holliday 模型的实验证据。他们使用含有 pMB9 质粒的大肠杆菌作为实验材料。每个大肠杆菌细胞通常含有约 20 个拷贝的 pMB9 质粒,为了增强质粒 DNA 之间发生同源重组的机会,他们使用氯霉素扩增的方法,将质粒的拷贝数增加到 1 000 个 / 细胞。

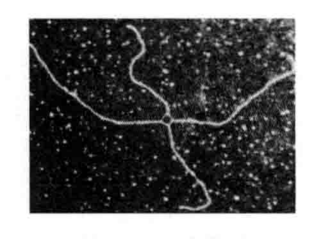


图 35-1　电镜下的 Holliday 结构

他们将质粒从细胞中抽取出来,并使用 CsCl 密度梯度离心纯化,最后放在电镜下观察,发现了呈 X 字形的结构(图 35-1)。但是,这种 X 字形结构并不意味着它一定是 Holliday 结构,因为双倍大小的质粒自身扭曲或者质粒的连环体在电镜下也会呈 X 字形。

为了将呈 X 状的 Holliday 结构与其他两种可能的结构区分开来,Potter 和 Dressler 使用限制性内切酶 *Eco*RⅠ预消化质粒,然后再放到电镜下观察。*Eco*RⅠ的消化将会破坏另外两种结构,只留下呈 X 状的 Holliday 结构。正如预期的一样,*Eco*RⅠ消化以后,留下了 0.5%~3% 呈 X 状的分子。进一步研究发现,这种呈 X 状的分子结构是对称的,其相反的 2 个臂的长度一样,并且具有相同的变性样式。当他们使用 Rec A 蛋白有缺陷的大肠杆菌突变株做同样实验的时候,却得不到呈 X 状的质粒分子。

Potter 和 Dressler 的实验结果为 Holliday 模型提供了一个直接的证据。正因为如此，虽然 Holliday 模型已被多次修改，但呈 X 状的 Holliday 结构依然是不可更改的。

（三）Aviemore 模型

虽然 Potter 和 Dressler 为 Holliday 模型提供了直接的证据，但 Holliday 模型仍然存在一些缺陷。例如，它并没有解释 2 个 DNA 分子的同源序列是如何配对以及单链切口是如何形成的，此外，它也不能解释在其他重组系统（如酵母）中得到的一些结果。于是，几个经改良的 Holliday 模型先后被提了出来，其中包括 Aviemore 模型或 Meselson-Radding 模型，这个新的模型有时也被称为单链断裂模型（single-strand break model）。

之所以将 Aviemore 模型称为单链断裂模型，是因为该模型认为（图 35-2），在 2 个配对的同源 DNA 分子中，只有 1 个 DNA 分子上的 1 条链产生切口，这条产生切口的链在被 DNA 聚合酶 I 催化的新链合成取代后，侵入到同源的 DNA 分子之中，至于 Holliday 结构的形成以及最后的分离，与原来的 Holliday 模型相比并没有做多少改动。

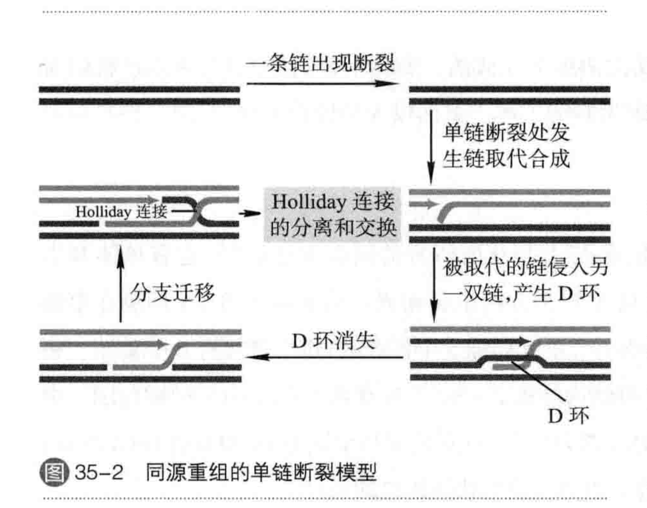


图 35-2 同源重组的单链断裂模型

（四）双链断裂模型（double-strand break model）

Aviemore 模型并不完美，于是又有人提出了双链断裂模型，该模型由 Szostak J. 等人于 1983 年共同提出，主要的实验证据来自酵母。与 Aviemore 模型不同的是，双链断裂模型认为，1 个 DNA 分子上两条链的断裂启动了链的交换，随后发生的 DNA 修复合成和切口的连接产生了 Holliday 结构，且具有 2 个半交叉点（half chiasma）。详细步骤如下（图 35-3）：①链断裂与切除，一个 DNA 分子发生双链断裂，

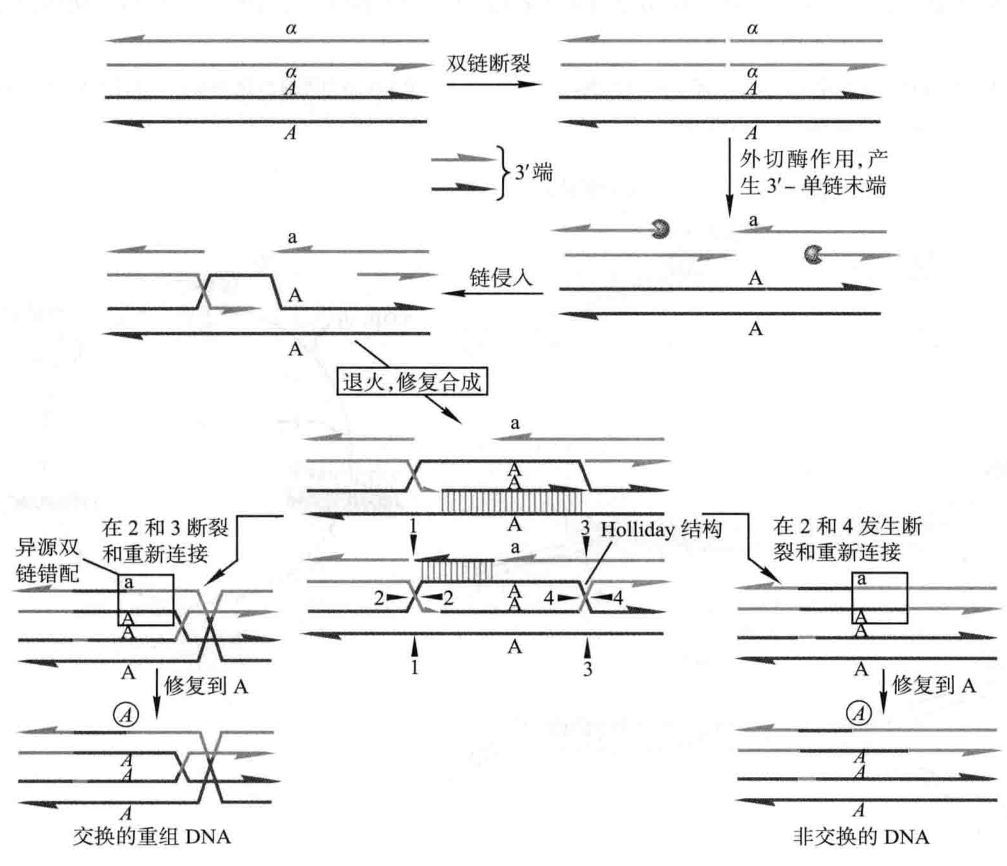


图 35-3 DNA 重组的双链断裂模型

随后在核酸酶的作用下，发生双链断裂的 DNA 分子上的两条链在双链裂口处发生降解。②链侵入，单链的一端在另一个双链 DNA 分子上寻找同源序列。如果发现同源序列，就侵入到另一个双链之中，进行链取代，从而形成异源双链。③退火与合成，在侵入的单链 3′端开始链的修补合成，同时，另一条单链的末端与被取代的链退火。④形成 Holliday 结构，DNA 合成和链取代导致第二个交换结构，即 Holliday 结构的形成。⑤ Holliday 结构的分离和交换。

二、参与同源重组的主要酶和蛋白质

无论哪一种重组都是在特定的蛋白质和酶的协助下完成的。实际上，对各种参与重组过程的蛋白质结构与功能的研究，大大促进了人们对重组机制的了解。下面以大肠杆菌为例，介绍一些与同源重组有关的蛋白质的结构与功能：

（一）RecA 蛋白

RecA 蛋白是同源重组中最重要的蛋白质，参与大肠杆菌所有的同源重组途径，它有单体和多聚体两种形式。多聚体由单体在单链 DNA 上从 5′→3′方向组装而成。多聚体的 RecA 环绕在单链 DNA 上形成一种有规则的丝状螺旋结构，平均每 1 个单体环绕 5 个核苷酸，每个螺旋有 6 个单体。每一个 RecA 单体由 3 个结构域组成，1 个大的中央结构域被 2 个位于 N 端和 C 端小的结构域包围。中央结构域负责与 DNA 和 ATP 结合，含有 2 个 DNA 结合位点，分别负责与单链 DNA 和双链 DNA 结合；N 端结构域与多聚体的形成有关；C 端结构域的功能是促进丝状结构之间的结合。

RecA 的主要功能包括：①促进 2 个 DNA 分子之间链的交换（图 35-4）；②作为共蛋白酶促进 LexA 阻遏蛋白的自我水解（参看第三十四章“DNA 的损伤、修复和突变”）。

要让 RecA 促进 2 个 DNA 分子之间的链交换，必须同时满足 3 个条件：① 2 个 DNA 分子中的 1 个必须含有单链区域，以便 RecA 能够结合上去；② 2 个 DNA 分子必须含有不低于 50 bp 的同源序列；③同源序列内必须含有 1 个自由的末端，以启动链的交换。

RecA 蛋白在同源重组的具体作用分为 3 个阶段：① RecA 的第一个 DNA 结合位点（初级位点）与

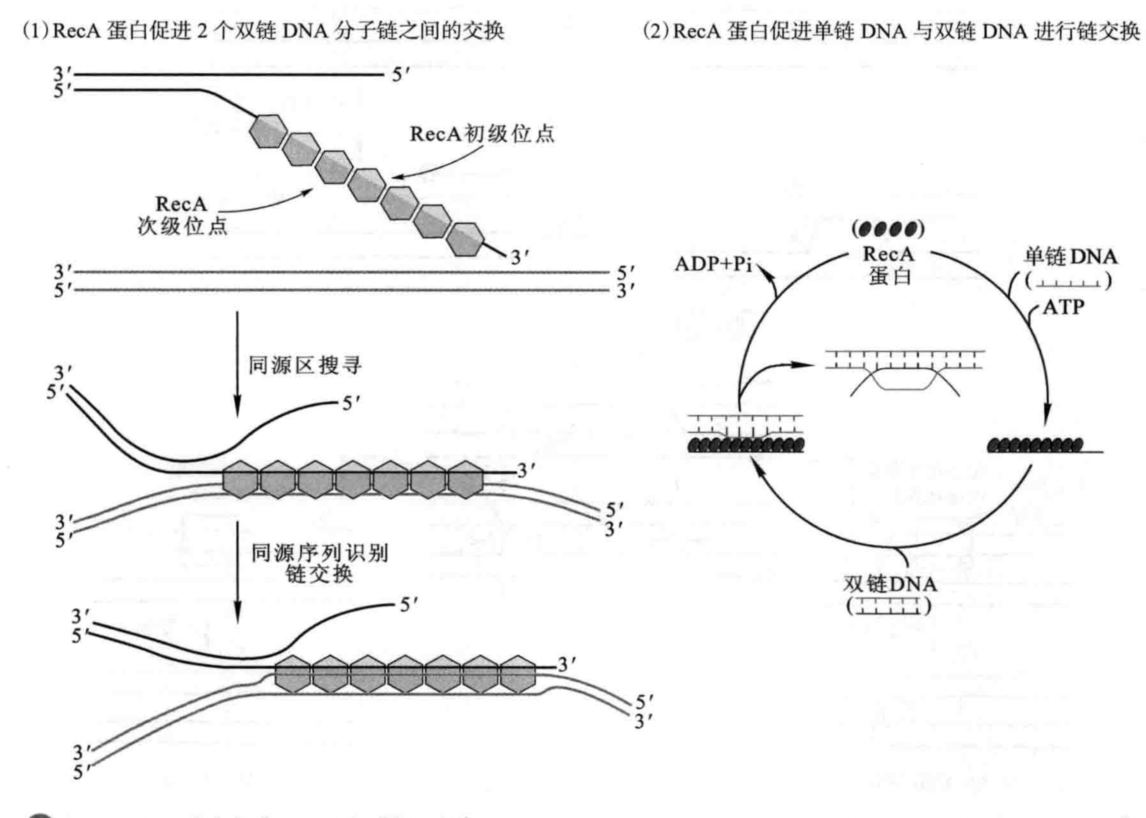


图 35-4　RecA 蛋白促进 DNA 之间进行链交换

Quiz1 用于基因工程的大肠杆菌通常都缺失编码 RecA 蛋白的基因，这是为什么？

单链 DNA 结合，包被 DNA，形成蛋白质 -DNA 丝状复合物；② RecA 的第二个 DNA 结合位点（次级位点）与 1 个双链 DNA 分子结合，形成三链 DNA 中间体，随后单链 DNA 侵入双链 DNA，寻找同源序列；③被 RecA 包被的单链 DNA 从 5′→3′方向取代双链 DNA 分子中的同源旧链，形成异源双链，并发生分叉迁移，在此阶段，由 RecA 驱动的链取代和分叉迁移需要 ATP 的结合。

在其他生物体中，也发现了 RecA 的类似物，例如古菌中的 RadA 和 RadB 蛋白，以及真核生物中的 Rad51、Rad57、Rad55 和 Dmc1 蛋白。

（二）RecBCD 蛋白

由 RecB、RecC 和 RecD 3 个亚基组成。RecBCD 蛋白具有 5 种不同的酶活性，分别是外切核酸酶Ⅴ、解链酶、核酸内切酶、ATP 酶和单链 DNA 外切酶活性。

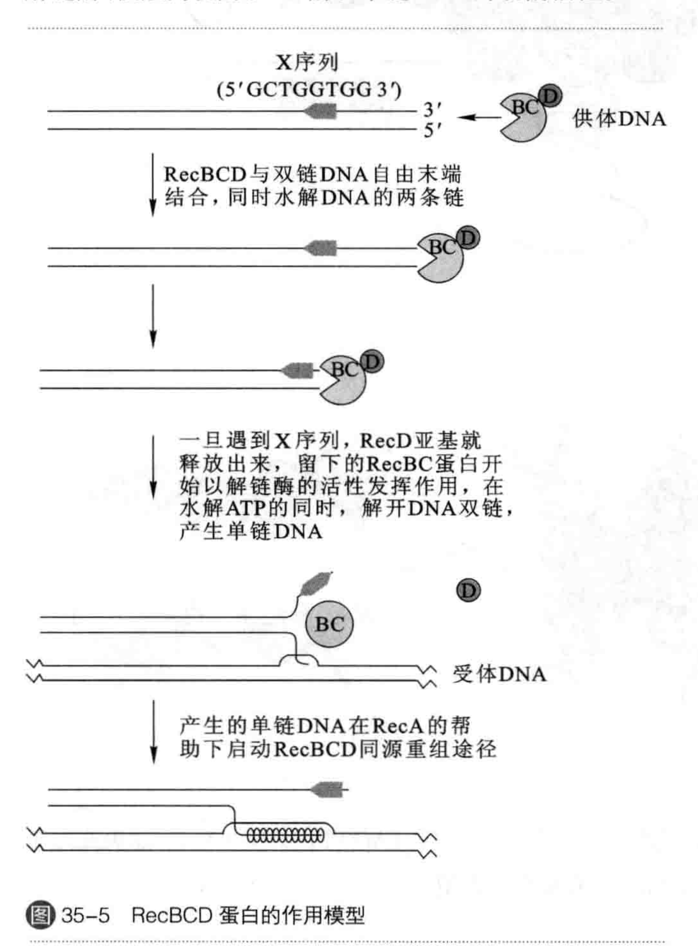


图 35-5　RecBCD 蛋白的作用模型

RecBCD 蛋白的功能是参与细胞内的 RecBCD 同源重组途径。在此途径中（图 35-5），RecBCD 首先与双链 DNA 分子自由末端结合，并依靠其外切核酸酶Ⅴ的活性，同时降解 DNA 的 2 条链。一旦它遇到 *chi*（X）序列，RecD 亚基就会从复合物中释放出来，留下来的 RecBC 蛋白开始以解链酶的活性发挥作用，在水解 ATP 的同时，解开 DNA 双链，产生单链 DNA，为 RecA 发挥作用铺平了道路，并最终启动链交换和重组反应。

X 序列是一段特殊的碱基序列，其一致序列是 5′-GCTGGTGG-3′，最早发现它与基因组上的重组“热点”有关联，即它的存在能提高重组的频率。显然，这与它能调节 RecBCD 的酶活性、刺激 RecBCD 重组途径有关。据估计，大肠杆菌基因组含有大于 1 000 个 X 序列。

（三）RuvA、RuvB 和 RuvC 蛋白

（1）RuvA　其功能是识别 Holliday 结构，协助 RuvB 蛋白催化分叉的迁移。它以一种特别的方式形成四聚体，呈四重对称，就像带有四片花瓣的花朵一样，特别适合与 Holliday 结构中的 4 个 DNA 双链区结合，从而促进分叉迁移过程中链的分离（图 35-6）。

（2）RuvB　此蛋白是一种解链酶，其功能是催化 Holliday 结构中分叉的迁移（图 35-6）。但它的作用需要 RuvA 的协助。与大多数解链酶一样，RuvB 也是一种六聚体蛋白，但与其他六聚体解链酶不同的是，RuvB 包被双链 DNA，而不是单链 DNA。电镜照片显示，RuvB 在溶液中是七聚体，但一旦与 DNA 结合，就转变为环状六聚体。电镜照片还显示，有 2 个 RuvB 六聚体与 RuvA 接触，各位于 RuvAB-Holliday 复合体的两侧。

（3）RuvC　此蛋白是一种特殊的核酸内切酶，负责催化 Holliday 结构的分离，因此被称为解离酶（resolvase）。它以对称的二聚体形式起作用，在 Holliday 结构的中央部位切开 4 条链中的 2 条，从而导致 Holliday 结构的拆分。由于 RuvC 同源二聚体与 Holliday 结构的结合是对称的，因此，RuvC 能以两种机会均等的方式与 Holliday 结构结合，致使 Holliday 结构能以两种机会相等的方式解离，但只有一种方式产生重组型 DNA（图 35-7）。

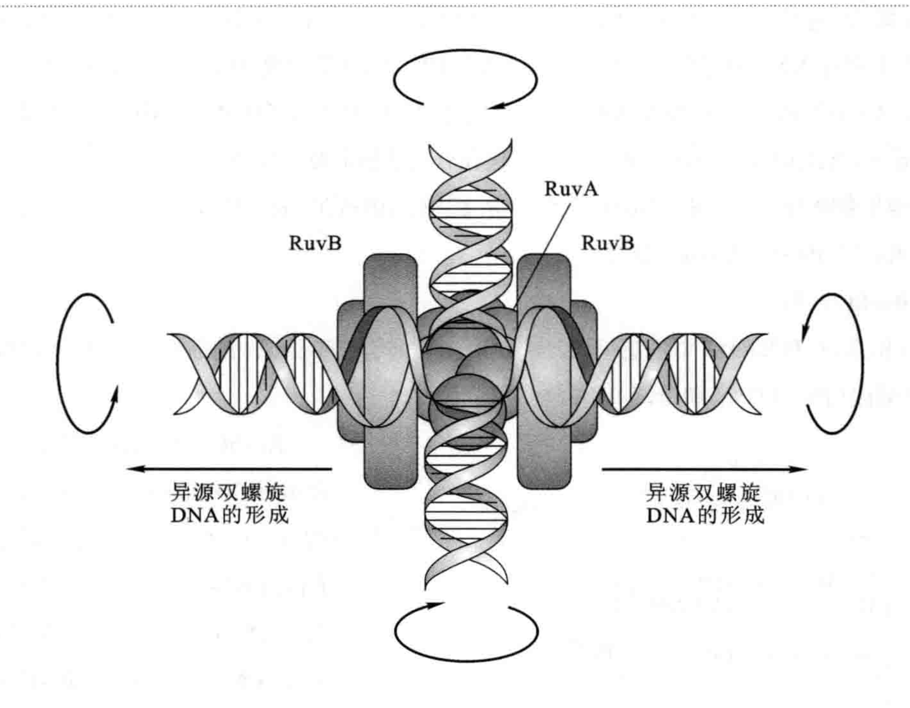


图 35-6 RuvA 和 RuvB 的作用模型

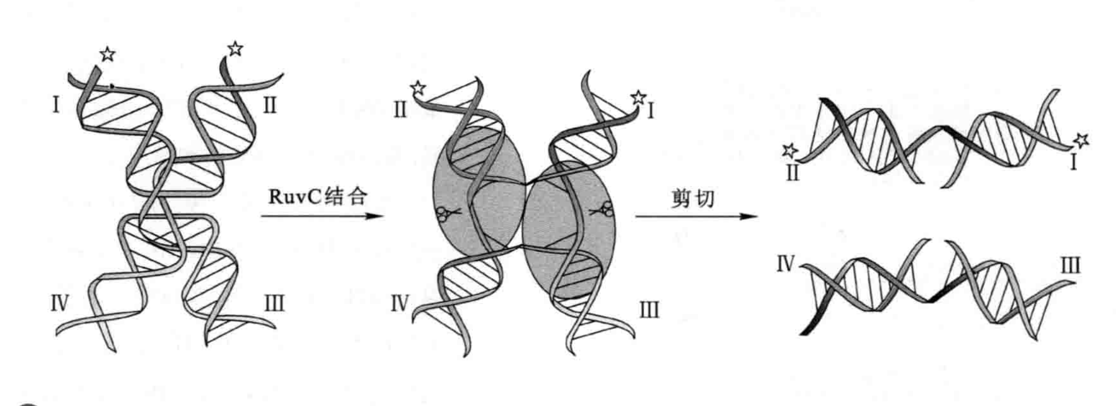


图 35-7 RuvC 的作用模型

RuvC 的作用具有一定的序列特异性，它作用的一致序列是 5′-(A/T) TT ↓ (G/C) -3′，箭头处为切点，只有当分支迁移到该一致序列的时候，RuvC 才有机会起作用。

三、细菌的同源重组

Quiz2 你认为 DNA 聚合酶 I 在 RecBCD 重组途径中起何作用，为什么不是 DNA 聚合酶 III？

e35-1 真核生物的同源重组

以大肠杆菌为例，细菌的同源重组最主要的是 RecBCD 途径(图 35-5)。在此途径中，除了 RecBCD 以外，还需要 RecA、SSB、RuvA、RuvB、RuvC、DNA 聚合酶 I 、DNA 连接酶和 DNA 旋转酶。此外，还需要 X 序列。

至于真核生物的同源重组，目前对人们对它的研究和认识还远远落后于原核生物，但从已经获得的数据看来，在不同的真核生物之间，同源重组的机制是高度保守的。

第二节 位点特异性重组

位点特异性重组是指在 DNA 特定位点上发生的重组，它几乎存在于所有的细胞中，需要专门的蛋白质识别特异性重组位点，并催化重组反应。虽然在很多情况下，它也需要在重组位点存在同源序列，但所需的同源序列并不长。与同源重组一样，位点特异性重组也发生链交换、形成 Holliday 结构、

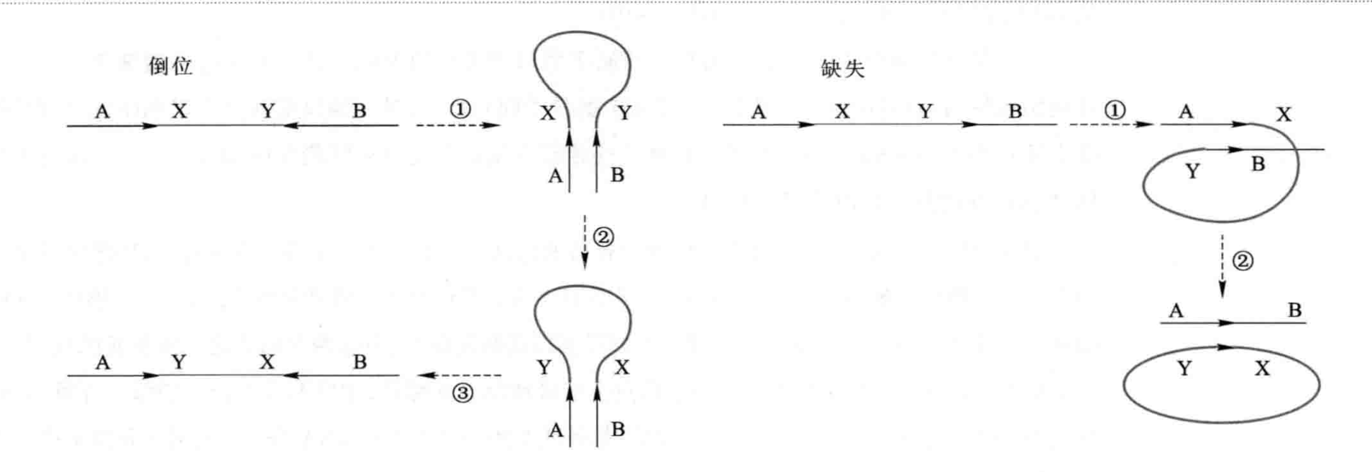


图 35-8 缺失性位点特异性重组和倒位式位点特异性重组图解

进行分叉迁移和 Holliday 结构解离的步骤。

位点特异性重组可以发生在 2 个不同的 DNA 分子之间，也可以发生在同 1 个 DNA 分子内（图 35-8）。如果是前者，通常就会导致 2 个 DNA 分子之间发生整合；如果是后者，就可能发生缺失或倒位（inversion）。缺失性位点特异性重组在 2 个重组位点上含有直接重复序列，而倒位式位点特异性重组在 2 个重组位点上具有反向重复序列（inverted repeat）。

位点特异性重组的主要功能包括：①调节噬菌体的整合；②调节基因表达。现分别举一例加以说明。

一、λ 噬菌体的位点特异性整合

当 λ 噬菌体感染大肠杆菌以后，其 DNA 首先通过 *cos* 位点自我环化。随后，噬菌体会在两条途径中做出选择：一条是裂解途径（lytic pathway），另一条是溶源途径（lysogenic pathway）。如果是裂解途径，噬菌体就会在较短的时间内大量复制，并裂解宿主细胞；如果是溶源途径，噬菌体就进入原噬菌体（prophage）的潜伏状态。处于这种状态的噬菌体几乎关闭所有基因的表达（参看第四十二章“原核生物的基因表达调控”）。

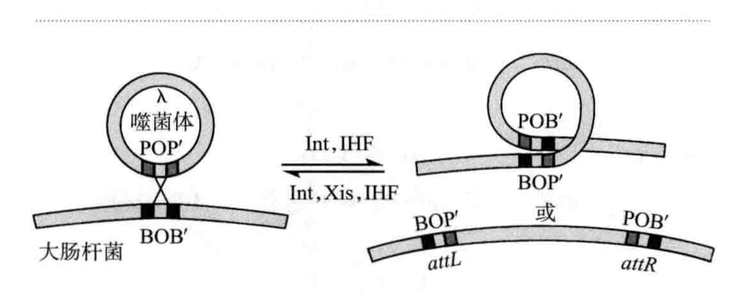


图 35-9 λ 噬菌体的位点特异性整合

要进入溶源途径，λ DNA 必须通过位点特异性重组的方式整合到大肠杆菌的染色体 DNA 上（图 35-9）。而大肠杆菌染色体本来就有专门的位点供其整合，此位点被称为附着位点 B（*attB*）。*attB* 只有 30 bp 长，中央含有 15 bp 的保守区域，重组反应就发生在该区域，该区域通常被称为 BOB′，其中 B 和 B′ 分别表示细菌 DNA 在这段保守序列两侧的臂。

噬菌体的重组位点称为 *attP*，其结构较为复杂，它的中央也含有与 *attB* 一样的 15 bp 保守序列，以 POP′ 表示，P 和 P′ 分别表示两侧的臂。*attP* 两翼的序列非常重要，因为含有一系列参与重组反应的蛋白质的结合位点。

λ 噬菌体整合需要 1 种自身编码的整合酶（integrase，Int）和 1 种宿主蛋白，即整合宿主因子（IHF）。两种蛋白质结合在 P 臂和 P′ 上，形成一种复合物，使 *attP* 和 *attB* 的 15 bp 序列能正确地并置在一起。Int 催化了重组过程的所有反应，包括一段 7 bp 长的分叉迁移。

重组的结果导致了整合的原噬菌体两侧成为两个附着点，它们的结构稍有不同，左边的 *attL* 结构

为 BOP′，右边的 *attR* 结构是 POB′（图 35-10）。

参与位点特异性重组反应的酶有一类属于酪氨酸重组酶家族。这一类重组酶通常含有 2 个保守的结构域，具有共同的反应机制，需要 4 个酶分子同时参与，链交换反应涉及在识别序列交错切开（2 个切点相距 6~8 bp）。所有的链切割和重连接反应实际上是系列转酯反应，不需要 ATP，这与Ⅰ型 DNA 拓扑异构酶十分相似（图 35-11）。

整个反应共分为 4 步：①有 2 个酶分子各自通过活性中心的 Tyr 残基上的羟基，亲核进攻识别序列上的 1 个磷酸二酯键，导致 2 个 DNA 分子各有 1 条链在识别序列内特定的位点被切开，形成 5′- 磷酸和 3′- 羟基，其中 5′- 磷酸与 Tyr 残基上的羟基以磷酸酯键相连；②两个切点之间发生转酯反应，即一个切点上的 3′- 羟基亲核进攻另一个切点上的磷酸酪氨酸酯键，重新形成 3′，5′- 磷酸二酯键，如此形成 Holliday 结构；③在短暂的分叉迁移后，另外两个酶分子在 2 个 DNA 分子上的另一条链上产生切口，反应同①；(4) 反应同②。

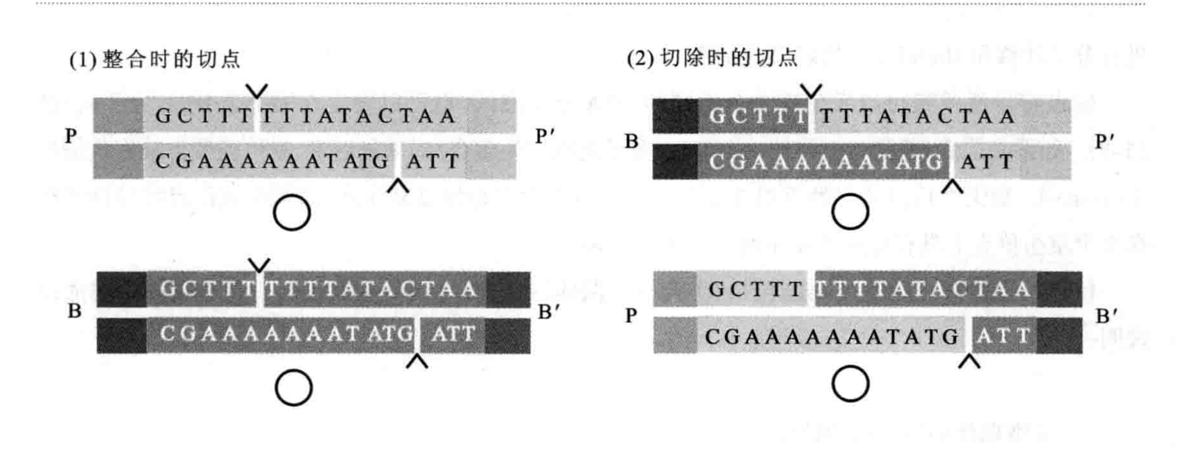


图 35-10 λ 噬菌体重组整合或切除时切点的序列

图 35-11 酪氨酸重组酶的作用机制（Topo 表示 DNA 拓扑异构酶，Int 表示整合酶）

λ噬菌体 DNA 从宿主染色体中的切除，除了需要 2 种自身编码的蛋白质 Int 和 Xis 以外，还需要几种细菌蛋白（IHF 和 Fis 蛋白）的参与。Xis 是一种切除酶（excisionase）。所有这些蛋白质都是与 *attL* 和 *attR* 上 P 臂和 P′臂结合，形成一种复合物，致使 *attL* 和 *attR* 上的 15 bp 保守序列正确地排列，从而有利于原噬菌体的释放。

Quiz3 λ 噬菌体 DNA 整合与切除都需要的蛋白质有哪些?

二、鼠伤寒沙门氏菌鞭毛抗原的转换

鼠伤寒沙门氏菌（*salmonella typhimurium*）的鞭毛由 H1 或 H2 鞭毛蛋白组成，但在一个特定的细胞内只有一种鞭毛蛋白表达。表达一种鞭毛蛋白的细胞偶然会转变为表达另外一种鞭毛蛋白的细胞（1/1 000），这种现象被称为相变（phase variation）。

相变的发生由位点特异性倒位（site-specific inversion）控制（图 35-12），并无遗传信息的丢弃，仅仅是通过倒位改变基因的方向，致使 H1 和 H2 只能表达一种。在一种方向，H2 操纵子能同时转录 H2 和 *rh1*，而 *rh1* 基因编码的是一种抑制 H1 转录的阻遏蛋白，于是 H2 表达，H1 就不能表达。然而，*hin* 基因编码的是重组酶，每隔一定时间它就催化位点特异性倒位，导致 H2 启动子离开 H2 操纵子。结果，H2 和 Rh1 均不能表达，H1 却能够表达。

Quiz4 若是 *rh1* 基因缺失，你认为鞭毛蛋白的表达会发生什么变化?

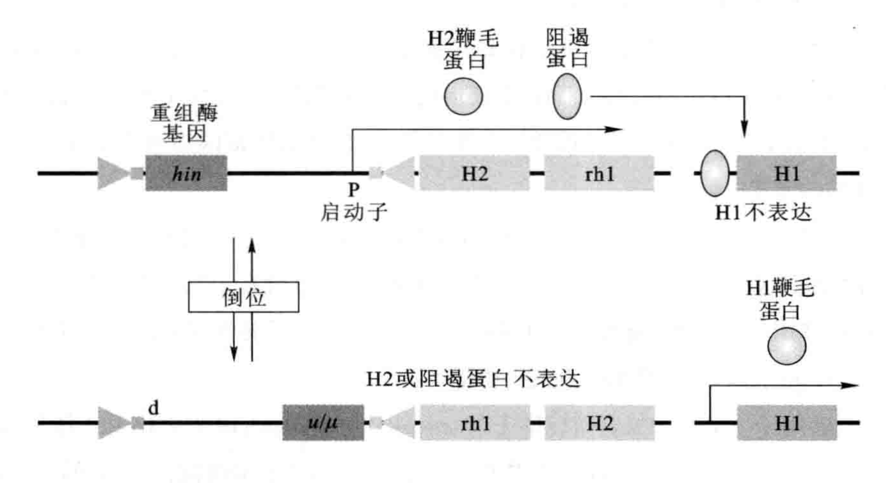


图 35-12 鼠伤寒沙门氏菌鞭毛抗原的转换

第三节 转座重组

转座重组是指 DNA 上的核苷酸序列从一个位置转位或跳跃到另外一个位置的现象，在原核生物和真核生物的基因组内均可以发生，发生转位或跳跃的核苷酸序列被称为转座子（transposon）。与同源重组和位点特异性重组不同的是，转座子的靶点与转座子并不存在序列的同源性。接受转座子的靶位点可能是随机的，也可能具有一定的倾向性。

最新的 DNA 序列分析表明，人、小鼠和水稻的基因组约有 40% 序列是由转座子衍生而来，而低等的真核生物和细菌内的比例较小，占 1%~5%。这说明转座子在从细菌到人类的基因和基因组进化过程中曾发挥过重要的作用。转座子的插入既可能导致基因的失活，也可能导致基因的激活，这与插入的位置有关。转座事件也可能导致基因组内的核苷酸序列发生转移、缺失、倒位或重复。

此外，转座子本身还可作为细胞内同源重组系统的底物，原因是在 1 个基因组内，2 个拷贝的同一种转座子提供了同源重组所需要的同源序列。两个转座子之间发生的同源重组又可导致缺失、插入、倒位和移位。

一、原核生物的转座

早在20世纪30年代后期，人们就在大肠杆菌内发现转座现象。首先被发现的转座子是插入序列(insertion sequence，IS)，它能够在DNA不同的位置之间进行转移，从一个位点插入到另一个位点，导致插入点(靶位点)的基因，以及和插入点基因在同一个操纵子内、但位于插入点基因下游的基因表达受阻。这种效应被称为极性效应(polar effect)。

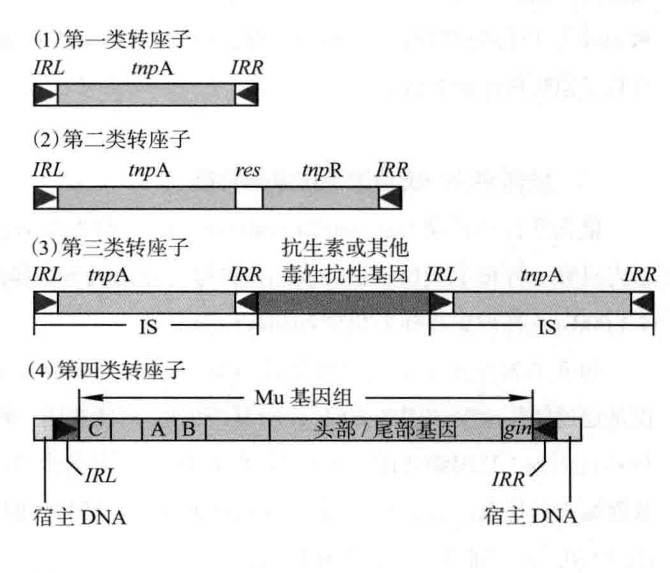


图35-13　细菌4类转座子的结构

到现在为止，已在细菌体内发现4类转座子(图35-13)，下面分别加以讨论。

(1) 第一类转座子　即插入序列，这一类转座子所具有的性质包括：①较小，长度一般在700~1 800 bp；②绝大多数两端含有10~40 bp长的反向重复序列(inverted repeat，IR)；③通常内部只有一个基因*tnp*A来编码催化转位反应的转座酶(transposase，Tnp)，而没有任何抗生素或其他毒性抗性基因；④通过“剪切”和“粘贴”的方式进行转座，转座结束后可导致插入点序列倍增；⑤有少数IS(如IS91)没有明显的IR，通过滚环复制与粘贴的方式进行转座。

(2) 第二类转座子　又称为复杂型转座子(complex transposon)，它们具有如下特征：①长度较长，为2.5~20 kb；②两侧含有较长的反向重复序列(35~40 bp)；③内部含有一个或几个结构基因，常见的结构基因有转座酶基因*tnp*A、解离酶基因*tnp*R和抗生素抗性基因；④转座完成以后可导致约5 bp长的靶位点序列倍增，从而在转座子两侧产生直接重复序列。

(3) 第三类转座子　又称为复合型转座子(composite transposon)，由2个IS和1段带有抗生素抗性或其他毒性抗性的间插序列组合而成，其中的2个IS位于转座子的两侧，相互间的方向可能相反，也可能一致。每一个IS带有转座酶，可以独立地转位，也可以与间插序列一起，作为一个整体进行集体转移。

(4) 第四类转座子　这一类转座子以大肠杆菌的Mu噬菌体(bacteriophage Mu)为代表。Mu噬菌体是一种温和性噬菌体，其DNA为线性双链，长度约为38 kb，两侧无反向重复序列，基因组有20多个基因，但只有*A*基因(编码转座酶)和*B*基因(与复制和转座有关)与转座有关联。在转座的过程中，可引起靶位点序列倍增。在溶源状态下，它能在宿主DNA的不同位置间随机转移。由于Mu转移的靶位点不固定，很容易诱发宿主细胞发生各种突变，其名称与它的诱变子(mutator)角色有关。

二、真核生物的转座

真核生物的转座现象最先由美国遗传学家Barbara McClintock在玉米中发现(详见本章科学故事)。现已明确，转座事件是发生在真核生物基因组内的普遍现象，各式各样的真核转座子已得以阐明。真核转座子与原核转座子有很多共同的性质，但也存在一些显著的差别，差别主要反映在转座的机制上，集中在两个方面：①真核生物转座过程中的剪切和插入是分开进行的；②很多转座子的复制通过RNA中间物来进行。

虽然目前还没有像原核生物那样的分类方法，但根据转座的机理，可将真核生物的转座子分为2类：第一类是逆转座子(retrotransposons)，其转座过程是DNA → RNA → DNA，需要RNA中间体；第二类是DNA转座子，其转座过程是DNA → DNA，无RNA中间体。

DNA 转座子又分为复制型 DNA 转座子(DNA transposon that transpose replicatively)和保留型 DNA 转座子(DNA transposon that transpose conservatively),其中,复制型在转位前后,原位置上的拷贝仍然存在,只是通过"复制"和"粘贴"的方式,复制一份到新的位点;而保留型在转座中,原来的拷贝被"剪切"后"粘贴"到新的位点,即原来的拷贝被原封不动地保留并转移到新的位点。

每一类转座子都有自主型(autonomous)和非自主型(non-autonomous)。自主型转座子的内部含有可读框,自身编码转座所必需的酶或蛋白质;非自主型转座子缺乏足够的编码能力,却保留了转座所必需的顺式元件,在合适的自主型转座子编码的转座酶的作用下,它们也可以进行转座。

(一) 复制型 DNA 转座子

这一类转座子通称为 Helitron,通过滚环复制的方式进行复制,然后再粘贴到新的位点。与其他大多数转座子不同的是:Helitron 的两端并没有反向重复序列,在转座以后,也不会导致靶位点序列倍增,但是,Helitron 的序列总是以 5′-TC 开始,3′-CTRR(R 表示嘌呤碱基)结束(图 35-14)。此外在 CTRR 序列的上游,有一段 16~20 nt 的回文序列可形成发夹结构,但回文序列本身并无保守性。Helitron 编码的蛋白质含有 5′→3′解链酶、核酸酶和连接酶的结构域,其英文名称部分来自于解链酶 Helicase 的前几个字母。

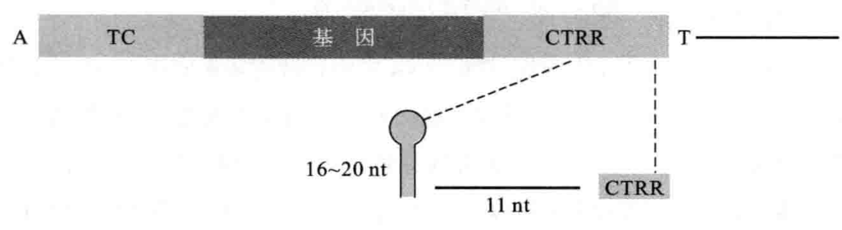


图 35-14 Helitron 转座子的结构

(二) 保留型 DNA 转座子

真核生物绝大多数 DNA 转座子属于这一类,例如玉米的 *Ac–Ds* 系统和果蝇的 *P* 元件(P element)。

(1) 玉米的 *Ac–Ds* 系统 McClintock 在玉米基因组发现的 *Ac–Ds* 系统是最有代表性的例子(图 35-15)。*Ac* 代表激活元件(activator element),属于自主型,带有全功能的转座酶基因;*Ds* 代表解离元件(dissociation element),属于非自主型,带有缺失的转座酶基因,由 *Ac* 突变而来。*Ac* 和 *Ds* 之间关系密切,一"主"一"仆",*Ds* 单独存在不能移位,因为它不能合成转座酶。同一个细胞内,必须有 *Ac* 的存在,*Ds* 才会移动。

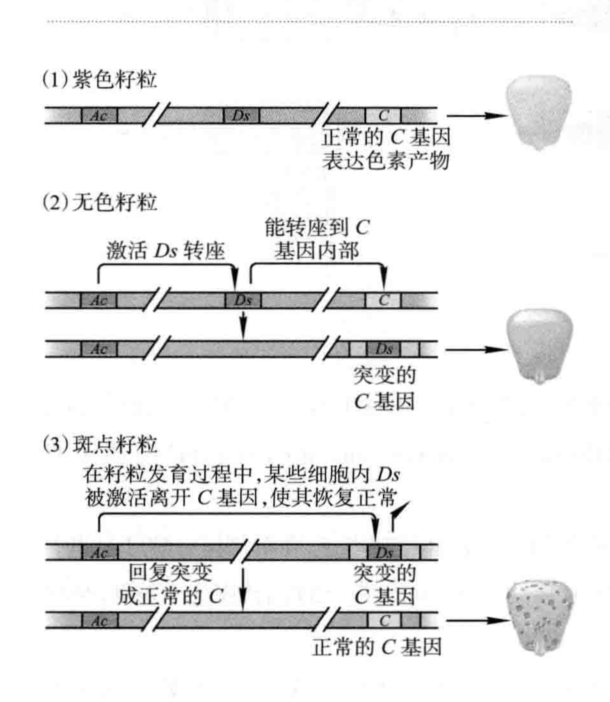


图 35-15 玉米的 *Ac–Ds* 系统

如果 *Ac* 或 *Ds* 插入到玉米种子的紫色色素基因 *C* 的内部,紫色色素基因就会被关闭,使玉米籽粒不能产生紫色色素而成为白色;如果 *Ds* 从 *C* 内部跳开,*C* 所受的抑制作用就会被解除,玉米籽粒又变成紫色。*Ac* 和 *Ds* 跳动得如此之快,使得受它们控制的颜色基因时关时开,于是一些玉米籽粒便出现了斑斑点点。

(2) 果蝇的 *P* 元件 完整的 *P* 元件长 2.9 Kb,两端含有 31 bp 的反向重复序列,内部含有 4 个可读框。由于它的转座酶基因的表达通常被抑制,因此,对果蝇造成的危害甚小。然而,当含有 *P* 元件的雄果蝇与缺乏 *P* 元件的雌果蝇交配的时候,生殖细胞内转座酶基因的表达就会被激活,从而

产生许多不能生育的突变体后代。

除了玉米的 *Ac–Ds* 系统和果蝇的 *P* 元件以外，人们还发现了“水手”元件(mariner element) 和 *MITE*(miniature inverted-repeat transposable element)。其中“水手”元件广泛存在于一系列生物体内，而 *MITE* 最先被发现在水稻、线虫体内。

e35-2 “水手”元件和MITE

(三) 逆转座子

逆转座子在真核生物的基因组中占很高的比例，它在结构、性质和转位的方式上与逆转录病毒的复制很相似(图 35–16)。首先是 DNA 转录为 RNA 中间物，然后 RNA 中间物被逆转录成双链 DNA，最后，双链 DNA 在新的位点整合到基因组 DNA 上。因此，在某种意义上，逆转座子可视为有缺陷的逆转录病毒。

带有逆转座子的供体DNA
逆转录酶基因
RNA
逆转录酶
cDNA
DNA
插入点
原来的逆转座子
插入到新位点上的逆转座子

图 35–16 逆转座子的转座机制

Quiz5 你如何设计一个实验，确定一个转座子是不是逆转座子?

根据两端的结构，逆转座子可进一步分为 LTR 逆转座子和非 LTR 逆转座子(图 35–17):LTR 逆转座子的两端是直接的长末端重复序列(long terminal repeat，LTR)；非 LTR 逆转座子没有 LTR 序列，但有一端终止于 1 小段重复序列(通常是 poly A)。无论是哪一种逆转座子，仍然有自主型和非自主型之分。如果是自主型的，其内部就含有 *gag* 基因和 *pol* 基因，但缺乏编码外壳蛋白的 *env* 基因。*pol* 基因能够编码蛋白酶、逆转录酶、核糖核酸酶 H 和整合酶。如果是非自主型的，内部序列大小会变化很大，缺乏大多数或者全部编码功能。

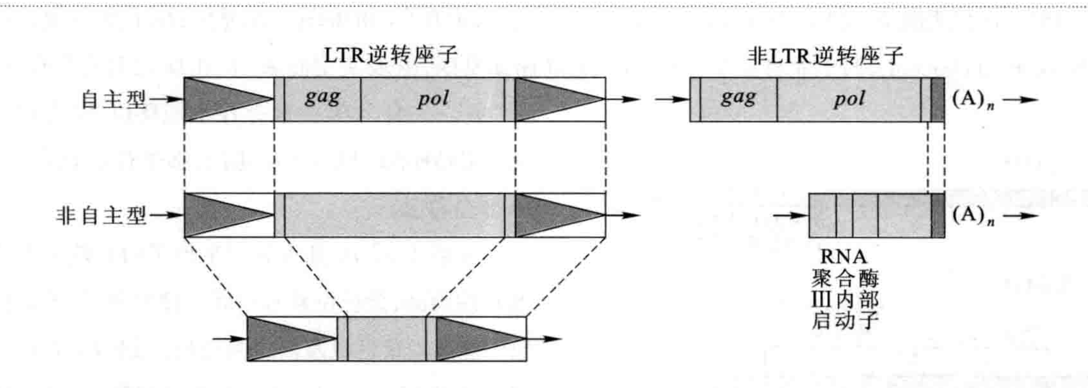


图 35–17 LTR 逆转座子和非 LTR 逆转座子的结构

果蝇基因组上的 *copia* 元件和酵母基因组上的 Ty 元件属于 LTR 逆转座子，长散布核元件(long interspersed nuclear element，LINE) 和短散布核元件(short interspersed nuclear element，SINE) 属于非 LTR 逆转座子。

(1) 酵母的 Ty 元件　Ty 元件是被研究的最为详尽的一种 LTR 逆转座子，在单倍体的酵母细胞中约有 35 个拷贝，散布在各个染色体 DNA 上，其 LTR 长度约为 330 bp。Ty 元件的转座效率很低，平均 10^4 世代才会发生 1 次。

Ty 元件可分为 6 大家族，但只有 Ty1、Ty2、Ty3 和 Ty917 具有转录的活性。它们转录出来的 mRNA 可占到单倍体酵母细胞总 mRNA 的 5% 以上。具有转录活性的 Ty 元件内部含有有活性的 *tyA*

（相当于 *gag*）基因和 *tyB*（相当于 *pol*）基因，分别编码一种 DNA 结合蛋白和逆转录酶，属于自主型的逆转座子。但由于缺乏编码外壳蛋白的 *env* 基因，因此，它们在细胞内并不能装配成具有感染性的病毒颗粒，但能够形成类似于病毒的颗粒。

（2）果蝇的 *copia* 元件　果蝇的 *copia* 元件与酵母的 Ty 元件结构相似，其长度约为 5.1 kb，每一个基因组有 20~60 个拷贝，散布在各染色体 DNA 上，其 LTR 的长度为 276 bp。每一个 *copia* 的内部含有单一的长可读框，编码一种多聚蛋白质，里面含有整合酶、逆转录酶和一种 DNA 结合蛋白，它们在果蝇细胞中也能形成类似于病毒的颗粒，但并无传染性。

copia 元件的转座效率为 10^3~10^4 世代发生一次，它在转座中可导致靶位点上的 5 bp 序列发生倍增。

（3）LINE　人类基因组并没有类似于 Ty 或 *copia* 元件的 LTR 逆转座子，但却发现了大量缺乏 LTR 的 LINE。LINE 散布在人类基因组 DNA 上，约占全部基因组 DNA 总量的 21%。LINE 也广泛存在于其他哺乳动物的基因组中。

人 LINE 的主要形式是 LINE-1（L1），其长度在 1~6 Kb，次要的形式为 LINE-2（L2）。完整的 L1 全长为 6.5 Kb，含有 2 个可读框，有一个相当于逆转录病毒的 *gag* 基因，编码一种 DNA 结合蛋白，另一个编码逆转录酶。

在人类基因组上，具有转录和翻译活性的完整的 LINE-1 并不多，大概有 50 多个，绝大多数都是长度不等的缺失性变体，已丧失了功能的基因。

L1-DNA 的转录终止不是很准确，转录的时候偶尔会多转录一些下游的碱基序列，有时又会提前终止而少转录一段序列。这两种转录产物经逆转录以后，产生的将是加长或截短的 L1 序列，这就是 L1 序列不均一的原因，其中被截短的 L1 很可能是没有功能的。有时，L1-RNA 翻译出来的酶活性偶尔将细胞内其他基因的 mRNA 逆转录成 cDNA，并将它们整合到基因组 DNA 上而产生无功能的假基因。通过这种手段产生的假基因既缺乏正常基因具有的内含子，又少了基因转录所必需的启动子和其他调控序列，因此不能表达。

（4）SINE　人类基因组含有一百多万个拷贝的 SINE，大概占基因组总量的 10%。它们散布在各个染色体 DNA 上，在靠近 5′端，含有 RNA 聚合酶Ⅲ所识别的内部启动子序列。绝大多数 SINE 属于

框 35-1　生化趣事——谁偷走了人的 DNA？

如果一个人细胞与一个细菌细胞进行“闪电约会”的话，它们绝不会相互交换各自的遗传信息。用更严格的科学术语来描述，就是在人和细菌之间不可能发生 DNA 的直接转移。然而，到了 2011 年 2 月，这种禁忌似乎已被打破。来自美国西北大学的研究人员首次发现，人类的一段 DNA 序列竟然已经跑到一些淋球菌（*Neisseria gonorrhoeae*）的基因组里了。淋球菌是导致人性病——淋病的罪魁祸首，其主要存在于生殖器官。这种人类历史上最古老的疾病之一在《圣经》中就有记载。据估计，全球有五千多万人染有淋病，治疗的方法是使用抗生素，但在过去的 40 多年中，许多抗生素已经陆续失效。

Mark T Anderson 在对临床上分离出来的淋球菌进行基因组序列分析的时候，在 15 例中分离到 3 例中含有一段与人类 L1-DNA 相同的序列。含有这种 L1-DNA 的淋球菌约占淋球菌分离物的 11%。而在与淋球菌亲缘关系十分密切的脑膜炎球菌（*Neisseria meningitidis*）中，并没有发现人类的 DNA。这说明，人类 DNA 转移到淋球菌是近期发生的事件。

一种细菌能从宿主中获取遗传信息，这在进化上的意义非同寻常。那些得到人类 DNA 的淋球菌可能会变成新的菌株。但这是否会给这些细菌提供某种好处尚不得而知。众所周知，不同的细菌之间，以及细菌和酵母之间可直接发生基因转移。现在发现人类的 DNA 被转移到细菌内，这听起来有点匪夷所思。至于其中的机制还不清楚。

Alu 家族或 Alu 序列。

Alu 家族具有以下特征:①通常含有限制性内切酶 *Alu* Ⅰ(识别序列为 AGCT)的切点,这是 Alu 序列名称的由来;②长度为 150~300 bp,属于高度重复序列;③散布于灵长类和鼠类的基因组 DNA 之中,人类的基因组中有 70 万 ~100 万的 Alu 序列,约占人类基因组全序列的 5%;④两端是 7~21 bp 的直接重复序列,长度随种类而异;⑤许多不同的 Alu 序列已得到分离和测序,结果表明约有 80% 的序列保守性;⑥ Alu 序列一般存在于基因之间和基因的内含子之中;⑦ Alu-RNA 不会翻译出蛋白质。

SINE 内部含有可被 RNA 聚合酶Ⅲ识别的启动子,因此可以在 RNA 聚合酶Ⅲ催化下进行转录,但并不编码逆转录酶和整合酶,故只能依赖于自主型的 L1 进行逆转录和整合。

三、转座的分子机制

由于转座酶是参与转座子转位的主要酶,因此,在详细阐述转座的分子机制之前有必要对几种转座酶的结构与功能做简单介绍。

到目前为止,已发现 5 类转座酶:① DDE- 转座酶,这一类转座酶含有高度保守的三联体氨基酸残基,即 D、D 和 E,这三个氨基酸残基是参与催化的 2 价金属离子(主要是 Mg^{2+})与酶分子的配位结合所必需的;② Y2- 转座酶,活性中心有 2 个 Y 残基参与催化;③ Y- 转座酶,活性中心只有 1 个 Y 残基参与催化;④ S- 转座酶,活性中心的 S 残基参与催化;⑤ RT/En- 转座酶,由逆转录酶和内切酶组合而成。这 5 类酶具有不同的催化机制来调节 DNA 链的断裂和重新连接。有的转座子使用"剪切"和"粘贴"的机制,有的则使用"复制"和"粘贴"的机制;有的使用 RNA 中间体,需要逆转录酶,有的则不经过 RNA 中间体;有的是"剪切"或"复制"与"粘贴"分开,有的则是同时进行。但一般可以将转座机制分为两种类型:一种是"剪切"和"粘贴"机制(cut-and-paste mechanism),另外一种是"复制"和"粘贴"机制(copy-and-paste mechanism)。

Quiz6 这些转座酶分子中的 D、E、Y 和 S 各代表什么氨基酸残基?

(1)"剪切"和"粘贴"机制　这种机制只是将起始位点上的转座子"剪切"下来,然后再"粘贴"到新的靶位点上去。显然,在转座完成以后,起始位点上的转座子序列已不复存在。参与此种转座机制的转座酶主要是 DDE- 转座酶。使用此种机制进行转位的转座子有:IS10、IS50、P 元件、Ac/Ds 元件和"水手"元件等。

(2)"复制"和"粘贴"机制　这种机制需要将起始位点上的转座子"复制"一份,然后再"粘贴"到靶位点上。显然,在转座完成以后,起始位点上的转座子序列依然存在。转座子的复制有两种方式,一种不需要 RNA 中间物,另一种则需要 RNA 中间物。

四、转座重组的调节

细胞内转座子的活性一般都维持在很低的水平,这是因为转座酶的活性受到严格的调控,以免伤害宿主细胞。调控可以在以下几个水平上进行:

(1) 转录水平　一般说来,转座酶的启动子是弱启动子,因此,转座酶的转录效率本来就低;其次,它的启动子通常有部分序列位于末端重复序列之中,这使得转座酶能够与自身的启动子序列结合进行自体调控,从而使转录的活性降低;再者,许多转座子内部含有抑制转录活性的 5- 甲基胞嘧啶;此外,某些转座酶还受到阻遏蛋白或反义 RNA 的负调控。

(2) 翻译水平　某些转座子 mRNA 翻译的起始信号隐蔽在特殊的二级结构之中,这使得它的起始密码子难以被核糖体识别,从而降低了它的翻译效率;此外,还可以通过翻译水平上的移框来减弱转座酶的翻译。

(3) 转座酶本身的稳定性　许多转座酶的稳定性很差,很容易被宿主细胞的蛋白酶降解,这在一定程度上降低了转座酶的活性。

(4) 转座酶活性的顺式调节　已发现某些转座酶对表达它的转座子或邻近的转座子的活性高,而

对其他位点上的同一种转座子的活性很低，这就限制了它对其他位点转座事件的影响。

(5) 宿主因素的影响　转座酶的活性经常受到宿主细胞内多种因子的调节，如 DNA 伴侣蛋白 (DNA chaperone)、IHF、HU 和 DnaA 蛋白等。

框 35-2　生化突破——睡美人转座子的复活

2009 年，国际分子和细胞生物学及生物技术方案和研究协会(ISMCBBPR)将睡美人(sleeping beauty, SB)转座子 / 转座酶 SB100X 评为当年的年度分子。此转座系统已经失活了一千多万年。1997 年，美国明尼苏达大学的 Zoltan Ivics 等人根据已积累的多种同家族转座子的系统发生数据，利用生物信息学的方法，对其进行基因重建，终于“唤醒”了其转座活性，所以有了这么美丽的名字。人们之所以选择它，是因为它能够在脊椎动物内实现可靠、稳定的基因转移。

我们已经知道，真核生物基因组序列的很大部分是由转座子衍生而来。以人基因组为例，约 45% 的序列来源于转座子。但是，绝大多数转座子已经没有活性。例如，从脊椎动物基因组分离出来的 Tc1/ 水手转座子超家族由于积累了大量突变，已经没有活性。鱼类基因组中的 TcE 就是其中的一个亚族代表。

Ivics 等人复活睡美人转座子的基本思路是：先重新激活转座酶，再重建能被激活的转座酶识别的顺式元件。他们从 8 种鱼中，共得到了 12 种鲑鱼类 TcE 的部分序列，然后从中确定可能具有功能的保守性蛋白质和 DNA 序列模体。在对其中突变了的转座酶的可读框进行概念翻译分析后，他们发现有 5 个区域在所有的 TcE 转座酶中是高度保守的：一个是靠近 N 端肽段的二元核定位信号，另一个是靠近肽链中央、功能不详的富含 Gly 模体，还有 3 段位于 C 端肽段，组成含有 DDE 的催化转位的结构域。多重序列比对显示转位酶编码区的突变随机性相当高。

他们在重新激活转座酶的过程中，先根据两个失活的亚特兰大三文鱼(atlantic salmon)的 TcEs 和一段彩虹鳟鱼 (rainbow trout) 的转座酶的序列，去除提前出现的翻译终止密码子和移框位置，一段段地拼凑出一个全长的可读框——SB3，然后进行活性测定，但发现没有活性。究其原因，是由于发生了非同义的核苷酸取代突变，SB3 多肽与转座酶的一致序列有 24 个位置不一样，其中 9 个在假定的功能结构域内。于是，他们在这个功能区进行了系统的逐一替换，最后终于得到了具有活性的转座酶。这种被复活的睡美人转座酶能在鱼类、小鼠和人类细胞中，以底物特异性的方式，结合在鲑鱼类转座子(salmonid transposon)两侧的 IR 上，通过“剪切 - 粘贴”的方式催化转位。

科学故事——“跳跃基因”的发现

1983 年的诺贝尔生理学或医学奖被授予美国遗传学家 Barbara McClintock。McClintock 得奖的时候已 81 岁高龄，离她提出“跳跃基因”假说的时间已近 32 年。她的假说长时间遭到冷遇甚至奚落，其前期境遇并不比遗传学的开山鼻祖 Mendel 好多少。Mendel 在修道院中孤独地进行豌豆杂交实验，通过数学的方法，整理出遗传的基本定律。在 1865 年，他就完成了这些实验，但一直等到 1900 年，他的研究结果同时被 3 位科学家在 3 个不同的国度再度发现后才被肯定。而 McClintock 花了近 50 年的时光在其狭窄的研究圈里，陪伴着她的是玉米。在那时候很多研究小组都以众取胜，更有大笔经费。但她却连一个研究助理也请不起。她的研究一直很少受到人们的注意。就在这种背景下，她发现了玉米有跳跃基因(*Ac* 和 *Ds*)，它们在染色体上可随情况不同跳来跳去，来去自如。这些基因可以控制或调节某些结构基因的活性。当她在 1951 年发表她的研究成果时，全世界大约只有寥寥数名遗传学家能了解她的研究报告。

其实在 20 世纪 30 年代，McClintock 对现代生物学就有了卓越贡献，她是第一个将观察及鉴别玉米染色体所需的细胞学技术发展成功的学者，从而证明了生物在形成配子时在染色体的互换中的确有遗传物质交换，为现代细胞遗传学的发展奠定了基础，而跳跃基因则是她后半生的另一高峰。

美国时代杂志曾经报导，McClintock 自幼就很有主见。她在 17 岁时进入 Cornell 大学就读，她本来想主修植物育种，但是由于该专业不收女生而改读植物系。1927 年她获得植物遗传学博士学位，由此与玉米开始了漫长的罗曼史。在 20 世纪 30 年代，女性科学家是没有地位的，因此她毕业以后像一个皮球一样，由一个机构滚到另一个机构，在 1942 年甚至失业。好在新成立的 Carnegie Institution of Washington 接受了她，并在 Cold Spring Harbor 遗传实验室给她一个研究员的职位。为了报答知遇之恩，她在那里一直工作到 1992 年病逝。

20 世纪 40 年代，McClintock 在 Cold Spring Harbor 发现了玉米基因中有一群控制玉米粒糊粉层颜色的突变基因，能从突变型回复成自然状态的野生型，而使子粒呈现红色斑点。这个结果跟她的同事 M. Rhoades 的发现十分类似，Rhoades 发现玉米粒的基因中有一对黑色的纯合子亲代理应都繁殖出黑色子代，居然得到不同颜色的子代，其比率为全黑：黑点：无色 = 12：3：1。

对于同样的结果，Rhoades 想到，玉米应有两个影响玉米粒颜色的基因，一个管颜色，另一管斑纹。当"斑点基因"出现时，无色的突变基因恢复成有颜色的形态，因此玉米粒的种皮显现斑纹。但他的学说必须有两个重要的前提，一是无色回复突变成有颜色的频率要很高，二是这种回复突变必须在有斑纹基因时方能发生。但是 McClintock 则不以为然，她假设除了颜色基因外另有两个基因，其一非常靠近颜色基因，可以直接控制其开或关；另一位于染色体上距离稍远处，可以控制第一个基因开关的速率。McClintock 将此二基因命名为控制元件，以别于结构基因。她的假说比发现细菌中有启动子和阻遏蛋白基因的 Jacob 和 Monod 足足早了十多年。

1956 年，McClintock 将有关的跳跃基因作了简要的阐述，她认为，她的实验结果已确定这些元件都存在于玉米的染色体上，她对这些控制元件的共同特征进行了总结：①可使特定的基因表达受到严格的调控，可开可关；②它们可在染色体上移动，也可在不同染色体间转位；③它们可以自动离去，使受控基因又回到本来面目；④它们可使染色体断裂，甚至引起染色体的移位、重复、缺失与倒位。

由于 McClintock 所持的"跳跃基因"观点并不被当时的科学界认可，因此她甚至放弃发表她的很多研究成果。她说："无人要读它，发表何用？"然而，时间是科学真理的试金石，到了 20 世纪 60 年代，分子遗传学家开始利用新的方法印证了她的发现。后来发展出来的基因克隆技术将 McClintock 在 30 年前发现的跳跃基因进行了详细的分析。这些导致她的假说最终被学术界接纳。

Ac–*Ds* 系统被研究得较为详细。*Ac* 是一种自主调节的基因，当它转移到某一基因附近就可调节该基因的活性。同时 *Ac* 也可使另一种被称为 *Ds* 的基因转移，结果引起染色体断裂及异常。*Ds* 可插入玉米的 *shrunken*（*Sh*）基因座次而引起 *Sh* → *sh* 之突变。位于 *Sh* 基因座次的基因编码的是蔗糖合成酶。1982 年，B. Burr 与 F. A. Burr 利用野生型玉米所制造的蔗糖合成酶 mRNA 以制备 cDNA，作为 *Sh* 基因座次的探针以寻找 *Ds* 插入的位置。再利用各种限制性内切酶将 *Ds* 元件切下来，并就较大的 DNA 加以分析。其结果显示，*sh* 突变型内的 *Ds* 基因长度为 20~22 kb。这些 *Ds* 元件插入 *Sh* 基因后就可引起突变。由此推论，若能移走此元件，就可以使突变型（*sh*）→野生型（*Sh*）。McClintock 曾分离出四种 *Ds* 诱发的突变型。经分析，两个突变型是在 *Sh* 基因座次靠近转录区插入了约 20 kb 的外源 DNA，而另外两种突变型则在 *Sh* 基因的 5′端插入。还有一种回复突变型在插入的 DNA 分子上发生了很大的重排，导致 *Sh* 基因恢复活性。

漫漫长夜后灿烂的黎明终于来临。在 1980 年后，名利如雪片般飞来，让 McClintock 应接不暇。1981 年，她荣获拉斯克基础医学研究奖（Albert Lasker Basic Medical Research Award）和以色列颁发的 Wolf Foundation 奖。同年被推选为芝加哥 Mac Arthur Foundation 的研究员，终其一生每年可获免税薪俸 6 万美元，而最高的奖项应该是 1983 年的诺贝尔奖。

McClintock 的事迹犹如一个科学界的神话故事，这位孤独的科学家长久在隐退中追寻真理，她的发现一向受到忽视，直至得到新技术和新手段的验证。而诺贝尔奖不过再度肯定罢了，虽然来得太迟了

一些。

为什么 McClintock 的成就要经过如此漫长的岁月才被科学界肯定呢？原因是这个观念对当时的遗传学家而言是完全陌生，而当她发表时，根本无人相信。英国爱丁堡大学遗传学教授芬肯说："有两个理由可以说明她的遭遇，第一，她触犯了当时生物学中'染色体是稳定的'信条。第二，她太急于将如此复杂的证据表现给同行，以致他们完全无法吸收。"芬肯记得，在 1948 年当他还是研究生时第一次遇见 McClintock，她立刻非常热切地向他解释这个发现，而且似乎逢人便如此。受到大家的漠视而得不到认同，McClintock 失望极了，从此她退隐且沉默，仿佛缩到蜗壳中，但仍不放弃研究。

e35-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Homologous_recombination（维基百科有关同源重组的内容）
2. http://en.wikipedia.org/wiki/Site-specific_recombination（维基百科有关位点特异性重组的内容）
3. http://bioweb.wku.edu/courses/biol22000/16Recombination/default.html（美国西肯塔基大学基础分子生物学课程有关重组的内容，内有小测验）
4. http://www.ncbi.nlm.nih.gov/books/NBK26845/（NCBI 提供的在线细胞分子生物学教材，有关重组部分的内容）

参考文献：

1. Anderson M T, Seifert H S. Opportunity and means: Horizontal gene transfer from the human host to a bacterial pathogen. *mBio*, 2(1): e00005-11, 2011.
2. Goodman M F, et al. Roles of DNA polymerase V and RecA protein in SOS damage-induced mutation. *Chem Rev*, 106(2): 406-419, 2006.
3. Lusetti S L, Cox M M. The bacterial RecA protein and the recombinational DNA repair of stalled replication forks. *Annu Rev Biochem*, 71: 71-100, 2002.
4. Pâques F, Haber J E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol Mol Biol Rev*, 63: 349-404, 1999.
5. Hallet B, Sherratt D J. Transposition and site-specific recombination: adapting DNA cut-and-paste mechanisms to a variety of genetic rearrangements. *FEMS Microbiol Rev*, 21: 157-178, 1997.
6. Kogoma T. Recombination by replication. *Cell*, 85: 625-627, 1996.

第三十六章　DNA 转录

DNA 是生物最重要的遗传物质,其贮存的遗传信息决定了细胞内外各种蛋白质的氨基酸序列。然而,在 DNA 分子中,以 A、G、T 和 C 4 种碱基编码的遗传信息并不能直接作为合成蛋白质的模板。按照 Francis Crick 于 1956 年提出的"中心法则"(central dogma)(图 36-1),DNA 首先作为模板指导 RNA 的生物合成,然后再由 RNA 直接指导蛋白质的生物合成。这种以 DNA 作为模板合成 RNA 的过程称为转录(transcription),而以 RNA 作为模板合成蛋白质的过程称为翻译(translation)。转录和翻译统称为基因表达(gene expression)。

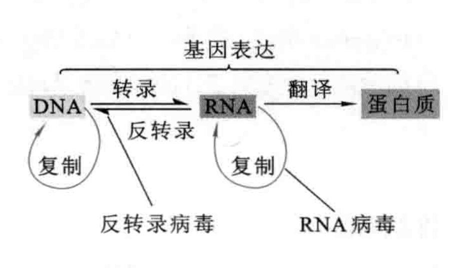


图 36-1　分子生物学的"中心法则"

细胞主要含有 3 种 RNA,即 rRNA,tRNA 和 mRNA,除此以外,还含有其他一些小 RNA,如真核细胞内的微 RNA、snRNA 和 7SL RNA。一个特定的细胞在特定的条件下所转录的所有 RNA 构成这个细胞的转录组(transcriptome)。转录组中所有的 RNA 分子都能与基因组 DNA 进行杂交,这表明它们都是以 DNA 为模板转录而来的。

本章将重点介绍 DNA 转录的一般特征、催化转录的 RNA 聚合酶的结构与功能以及转录的详细机制。

第一节　DNA 转录的一般特征

DNA 转录可发生在原核细胞的细胞质、真核细胞的细胞核、线粒体和叶绿体的基质中,这四大转录系统具有以下共同的特征:

(1) 转录具有选择性和不对称性,只发生在 DNA 分子上某些特定的区域。

转录与 DNA 复制的一个显著差别是转录只发生在 DNA 分子上特定的区域。对于一个 DNA 分子而言,并不是所有的序列都会转录,有的序列从来不会转录,而能转录的序列也不是始终都在转录。此外,DNA 也并非两条链都会转录。事实上,某些基因以 DNA 的一条链为模板,而其他基因以另一条链为模板。对某一个特定的基因来说,双螺旋 DNA 分子上作为模板的那一条链称为模板链(template strand),与模板链互补的那一条链称为编码链(coding strand)。模板链又名无意义链(nonsense strand)或 Watson 链,编码链又名有意义链(sense strand)或 Crick 链。

Quiz1　DNA 复制的时候,能够直接插入少量 U。你认为,DNA 转录的时候,能否直接插入少量的 T?为什么?

(2) 以四种 NTP——ATP、GTP、CTP 和 UTP 为前体,还需要 Mg^{2+}。

(3) 转录需要模板、解链,但不需要引物。

(4) 最先转录出的核苷酸通常是嘌呤核苷酸(约占 90%)。

(5) 转录的方向总是从 5′→ 3′。

(6) 转录也具有高度的忠实性,但比 DNA 复制低。

转录的忠实性是指一个特定的基因转录具有相对固定的起点和终点,而且转录过程基本遵守碱基互补配对原则。然而,转录的忠实性明显要低于 DNA 复制,原因主要是 RNA 聚合酶没有 3′- 外切酶活性,而且,机体在一定程度上可以容忍转录相对低的忠实性。

Quiz2　为什么机体可以容忍转录相对低的忠实性?

(7) 有高度的进行性。

正常的转录一旦进入延伸，就不会中途终止，这一点有别于 DNA 复制。一旦 RNA 聚合酶在转录的时候从模板链解离，则解离下来的酶不可能与解离点重新结合。

(8) 转录受到严格调控。

调控的位点主要是在转录的起始阶段，也可在转录的终止阶段。

第二节 依赖 DNA 的 RNA 聚合酶

DNA 转录是一种复杂的酶促反应，起主导作用的是 RNA 聚合酶，RNA 聚合酶全称为依赖于 DNA 的 RNA 聚合酶(DNA-dependent RNA polymerase，RNAP)。然而，生物化学家最先得到的催化 RNA 合成的酶是多聚核苷酸磷酸化酶(polynucleotide phosphorylase，PNP)。但 PNP 的一些性质使其不适合催化基因的转录，例如它不需要模板，使用 NDP 代替 NTP，合成的 RNA 序列取决于 NDP 的种类和相对浓度，这些性质无法保证基因转录的忠实性。后来才知道，PNP 在细胞内的功能是降解而不是合成 RNA。真正催化 DNA 转录的酶直到 1960 年才从大肠杆菌中得到。

RNA 聚合酶所催化的反应通式为：$n\text{NTPs} \xrightarrow{\text{DNA 模板},Mg^{2+}} (\text{NMP})_n + n\text{PPi}$

反应的机理与 DNA 聚合酶极为相似，都以 DNA 为模板，并从 5′→3′方向催化多聚核苷酸的合成，并涉及 2 个 Mg^{2+} 的双金属催化(见本章第三节“细菌的 DNA 转录”)，主要产物是与 DNA 模板链序列互补的 RNA。上述反应原则上是可逆的，但细胞内的焦磷酸酶可将产生的 PPi 迅速地水解掉，从而使得反应平衡向右倾斜。

虽然 RNA 聚合酶与 DNA 聚合酶在催化机制上十分相似，但这两类聚合酶仍有许多差别，总结起来包括：① RNA 聚合酶只有 5′→3′的聚合酶活性，无 5′-外切酶或 3′-外切酶的活性，RNA 聚合酶缺乏 3′-外切酶的活性是转录忠实性不如复制的主要原因；② RNA 聚合酶本身就能够促进 DNA 双链解链；③ RNA 聚合酶催化转录的从头合成，不需要引物；④ RNA 聚合酶与进入的 NTP 上的 2′-OH 有多重接触位点，借此区分 NTP 和 dNTP；⑤ RNA 聚合酶在转录的起始阶段，DNA 分子会形成褶皱(DNA scrunching)，以保证在发生无效转录(abortive transcription)时，仍然能与启动子结合；⑥在转录过程中，转录物与模板不断解离，而在复制中，DNA 聚合酶上开放的裂缝允许 DNA 双链从酶分子上伸展出来；⑦ RNA 聚合酶在转录的起始阶段受到多种调节蛋白的调控；⑧ RNA 聚合酶的底物是 NTP，而不是 dNTP，并由 UTP 代替 dTTP；⑨ RNA 聚合酶启动转录需要识别启动子；⑩ RNA 聚合酶的反应速率低，平均值只有 50 nt/s。

一、细菌的 RNA 聚合酶

细菌除了较短的 RNA 引物是由引发酶催化合成以外，细胞内其他所有的 RNA 都是由同一种 RNA 聚合酶负责转录的。

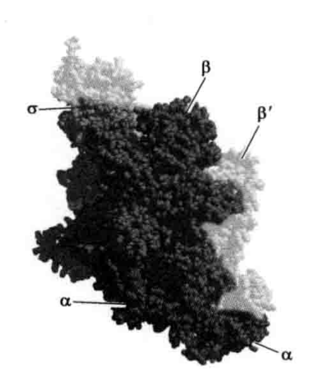
(一) 大肠杆菌 RNA 聚合酶的组成

以大肠杆菌为例，细菌的 RNA 聚合酶有核心酶(core enzyme)和全酶(holoenzyme)两种形式。核心酶的组成包括 2 个 α 亚基、1 个 β 亚基、1 个 β′亚基和 1 个 ω 亚基($\alpha_2\beta\beta'\omega$)(表 36-1)，其中 β′亚基结合有 2 个 Zn^{2+}。全酶则由核心酶和 σ 因子组装而成。σ 因子是可变的，如大肠杆菌一共有 7 种——σ^{70}、σ^{54}、σ^{32}、σ^{38}、σ^{24}、σ^{28} 和 σ^{19}。其中，σ^{70} 最常见，参与大肠杆菌在正常的条件下维持细胞生存必需的管家基因的转录，如编码糖酵解、三羧酸循环和 DNA 复制中各种酶或蛋白质的基因，而其他 σ 因子在大肠杆菌处于特殊的环境条件(如热休克和氮饥饿)时起作用。

► 表 36-1 大肠杆菌 RNA 聚合酶全酶的组成及其功能分工

亚基	基因	*Mr*/kb	数目 / 酶	功能
α	*RopA*	36	2	核心酶的组装,转录起始,与调节蛋白的作用
β	*RopB*	151	1	转录的起始和延伸
β′	*RopC*	155	1	与 DNA 非特异性结合
ω	*RopZ*	11	1	促进核心酶的组装,为 β′ 亚基的分子伴侣,在体外为变性的 RNA 聚合酶成功复性所必需
σ^{70}	*RopD*	70	1	启动子的识别

Quiz3 许多与核酸结合的蛋白质通过特殊的通道与目标核酸结合,而通道壁一般都富含碱性氨基酸。为什么?



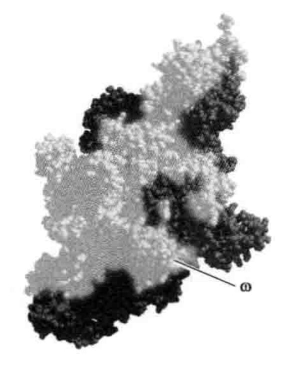


图 36-2 *Taq* RNA 聚合酶核心酶的三维结构两面观(Nelson 等,2008)

(二) 细菌 RNA 聚合酶的三维结构

细菌的 RNA 聚合酶的各亚基组装成类似于 DNA 聚合酶的“手掌”状结构。以水生嗜热菌的 *Taq* RNA 聚合酶为例(图 36-2),其两个 α 亚基分别与 β 亚基和 β′ 亚基的一端结合,ω 亚基环绕在 β′ 亚基的 C 端,而 β 亚基和 β′ 亚基仿佛是构成螃蟹钳子的一对螯,酶的活性中心正好位于两只螯之间长约 5.5 nm、直径约 2.7 nm 的通道内。与酶外表面大多带负电荷的性质相反,通道壁上却带有正电荷,适于结合 DNA 和 RNA,在底部始终结合 1 个起催化作用的 Mg^{2+}。如果有 σ 因子结合,由它跨越核心酶的顶部,导致两鳌收拢,通道因此变窄了近 1 nm。

与其他任何一种 RNA 聚合酶一样,细菌的 RNA 聚合酶除了含有结合 DNA 模板的通道以外,还含有结合单链 DNA 的通道、NTP/Mg^{2+} 进入活性中心的通道以及 RNA 离开通道。

(三) 细菌 RNA 聚合酶的特异性抑制剂

细菌 RNA 聚合酶特异性的抑制剂包括利福霉素(rifamycin)和利链霉素(streptolydigin),它们都能与细菌的 RNA 聚合酶形成紧密的复合物,作用的对象都是 β 亚基,结合的部位都不在活性中心,但抑制的原理完全不同,前者能阻断长度仅为 2~3 nt 的新生转录物从 RNA 聚合酶离开通道的离开,从而抑制转录的起始,后者则能阻止 RNA 酶在催化过程中必须经历的构象变化,从而抑制转录的延伸。它们并不抑制古菌和真核生物的核 RNA 聚合酶。正因为如此,两者均可被用作抗生素。

二、真核生物的 RNA 聚合酶

(一) 真核生物 RNA 聚合酶的分类与组成

真核生物的 RNA 聚合酶出现了功能分工,不同性质的 RNA 由不同类型的 RNA 聚合酶催化转录(表 36-2),在细胞核至少有 3 种 RNA 聚合酶,即 RNA 聚合酶 Ⅰ、Ⅱ 和 Ⅲ,三者也分别被称为 RNA 聚合酶 A、B 和 C。其中,RNA 聚合酶 Ⅰ 催化核仁内的 rRNA(5S rRNA 除外)的合成,但对于锥体虫(*Trypanosome*)而言,其可变的表面糖蛋白(variable surface glycoprotein,VSG)的基因转录也由聚合酶 Ⅰ 催化;RNA 聚合酶 Ⅱ 主要负责催化 mRNA 或者以 mRNA 为前体的 RNA(如 miRNA)的合成,同时也催化细胞内含有帽子结构的 snRNA 和 snoRNA 的合成;RNA 聚合酶 Ⅲ 催化细胞内各种较为稳定的小 RNA 的合成,例如 tRNA、5SrRNA、7SL RNA 和无帽子结构的 snRNA 等。

除了上述三种 RNA 聚合酶以外,一些植物的细胞核还含有另外两种 RNA 聚合酶,即 RNA 聚合酶 Ⅳ 和 Ⅴ,其中 RNA 聚合酶 Ⅳ 催化 siRNA 的合成,RNA 聚合酶 Ⅴ 催化促进染色质浓缩的 siRNA 的合成。此外,线粒体和叶绿体也含有 RNA 聚合酶,其中线粒体内的 RNA 聚合酶由核基因编码,只由一个亚基组成,负责线粒体 DNA 上所有基因的转录,它在结构上并不像细菌的 RNA 聚合酶,而是与 T7 噬菌体 RNA 聚合酶相似。而叶绿体有两种不同的 RNA 聚合酶,一种由核基因组编码,在结构上类似于线粒体内的 RNA 聚合酶,负责催化叶绿体内 rRNA 或 tRNA 等管家基因的转录;而另一种由叶绿体基因组编码,类似于细菌的 RNA 聚合酶,但对利福霉素不敏感,负责催化光合链上的复合体中由叶绿体基因组编码的各亚基基因的转录。

真核细胞核的三种聚合酶都是庞大的多亚基蛋白(2 个大亚基和 12~15 个小亚基),如酵母的

► 表 36-2 真核生物 5 种 RNA 聚合酶结构与功能的比较

名称	细胞中的定位	组成	对 α 鹅膏蕈碱的敏感性	对放线菌素 D 的敏感性	转录因子	功能
RNA 聚合酶 Ⅰ(A)	核仁	多个亚基组成	不敏感	非常敏感	1~3 种	rRNA 的合成(除了 5S rRNA)
RNA 聚合酶 Ⅱ(B)	核质	多个亚基组成	高度敏感(10^{-9}~10^{-8} mol/L)	轻度敏感	8 种以上	mRNA、miRNA、snoRNA、lincRNA、具有帽子结构的 snRNA 的合成
RNA 聚合酶 Ⅲ(C)	核质	多个亚基组成	中度敏感	轻度敏感	4 种以上	小 RNA,包括 tRNA、5S rRNA、无帽子结构的 U6-snRNA、7SL RNA、7SK RNA、RMP RNA、端粒酶 RNA、少数 miRNA 和 lincRNA、某些病毒的 RNA 等合成
线粒体 RNA 聚合酶	线粒体基质	单体酶	不敏感	敏感	2 种	所有线粒体 RNA 的合成
叶绿体 RNA 聚合酶	叶绿体基质	有两种	不敏感	敏感	3 种以上	所有叶绿体 RNA 的合成

RNA 聚合酶 Ⅱ 共有 12 个亚基。总的相对分子质量在(5~7)×10^5,有几个亚基是相同的。每一种聚合酶都含有由 10 个亚基组成的极为保守的核心结构,其中 2 个最大亚基的序列与细菌 RNA 聚合酶 β 和 β′ 亚基相似。此外,它们都还含有细菌 RNA 聚合酶 α 和 ω 亚基的同源物,但无任何亚基与细菌的 σ 因子同源。

e36-1 酵母 RNA 聚合酶Ⅱ

Quiz4 为什么真核 RNA 聚合酶没有任何亚基与细菌的 σ 因子同源?

需要特别注意的是,所有的真核生物的 RNA 聚合酶 Ⅱ 的最大亚基,即与细菌 β′ 亚基同源的亚基在 C 端都具有一种特殊的结构域,这种羧基端结构域(carboxyl-terminal domain,CTD)含有多个拷贝的呈高度保守的七肽重复序列,其一致序列是富含羟基氨基酸的 YSPTSPS,因此可被磷酸化修饰。在酵母细胞中,该序列重复了 26 次,哺乳动物则重复了 56 次。这种重复序列是 RNA 聚合酶 Ⅱ 的活性所必需的。缺失突变实验表明,酵母的 CTD 至少需要 13 段重复序列才能保证酵母细胞的生存。其他研究还表明,CTD 没被磷酸化的 RNA 聚合酶 Ⅱ 参与转录的起始,而 CTD 被磷酸化以后,转录即从起始阶段进入延伸阶段。高度磷酸化的 CTD 由于相邻带同种电荷的磷酸基团之间的排斥而突出暴露在外,远离酶的球状核心区达 50 nm,这为参与转录以及转录后加工的各种蛋白质因子对它的识别和结合提供了绝佳的舞台。事实证明,磷酸化的 CTD 对于 mRNA 转录的后加工是必需的。

Quiz5 你认为这段序列能够形成 α 螺旋或 β 折叠吗?为什么?

真核生物的 RNA 聚合酶在三维结构上与细菌的十分相似,不但分子的整个形状相似,而且各同源亚基在空间上总的排布也相同(图 36-3)。但与细菌的 RNA 聚合酶不同的是,真核生物 RNA 聚合酶本身并不能直接识别启动子,而必须借助于转录因子才能有效地与启动子结合,并启动转录。

(二) 真核细胞核 RNA 聚合酶的抑制剂

真核生物三种细胞核 RNA 聚合酶不受利福霉素或利链霉素的抑制,但是,它们对来源于某些毒蘑菇,例如白毒伞(*Amanita Phalloides*)体内的一种环状八肽毒素——*α*- 鹅膏蕈碱(图 36-4)的抑制表现出不同程度的敏感性:RNA 聚合酶 Ⅱ 最为敏感,10^{-8}~10^{-9} mol/L 的 *α*- 鹅膏蕈碱就能完全抑制其活性,其次是 RNA 聚合酶 Ⅲ,而 RNA 聚合酶 Ⅰ 则不敏感。*α*- 鹅膏蕈碱的作用机制是通过与聚合酶分子上一段特殊的桥螺旋(bridge helix)的结合,阻止聚合酶的移位,致使转录延伸受阻。由于 *α*- 鹅膏蕈碱与 RNA 聚合酶 Ⅱ 大亚基形成的是紧密的 1∶1 复合物(K_d=10^{-9} mol/L),与聚合酶 Ⅲ 形成的是比较松散的复合物(K_d=10^{-6} mol/L),而与 RNA 聚合酶 Ⅰ 基本不结合,因此这三种聚合酶对它的敏感性才会各不相同。

Quiz6 你认为哪一类生物的 RNA 聚合酶Ⅱ不受 α- 鹅膏蕈碱的抑制?为什么?

Quiz7 如何确定真核生物的一个基因由哪一种酶催化转录?

放线菌素 D(actinomycin D)能够插入到 GC 碱基对之间,致使 DNA 双螺旋的小沟变宽和扭曲,从

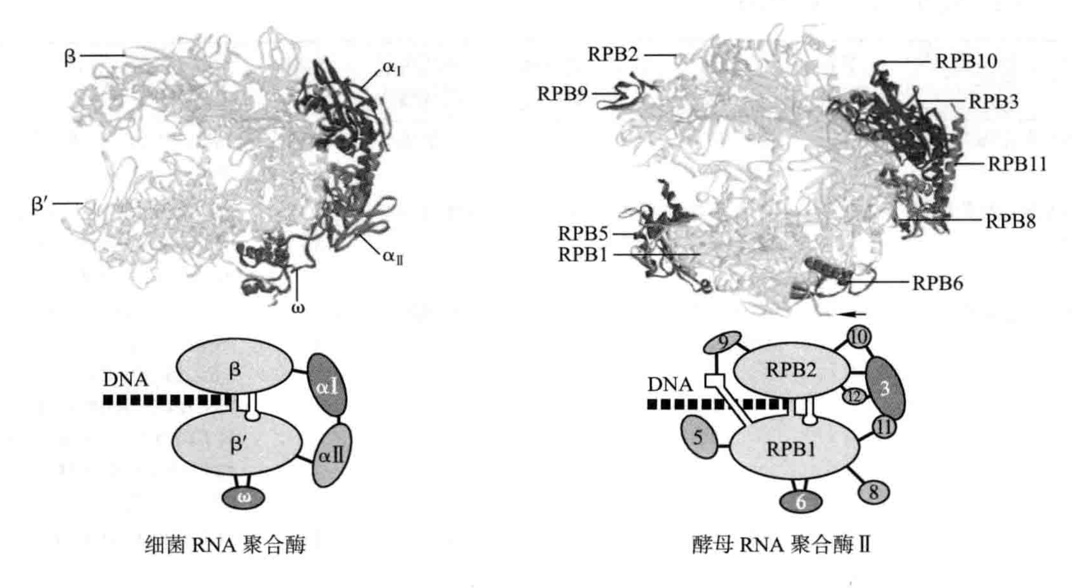


图 36-3　细菌生物的 RNA 聚合酶在三维结构上的比较（Lodish 等，2007）

框 36-1　身边的生物化学——路边的“野蘑菇”你不要采！

除非你是研究蘑菇的专家，否则就不要随意采集路边免费的蘑菇去食用，原因是你采到的蘑菇很可能是有毒的。尽管在民间有很多所谓鉴别有毒和无毒蘑菇的方法，但基本上都无科学依据。有的毒蘑菇和无毒蘑菇看起来非常像，有时专家也需要通过显微镜才能区分开来。



白毒伞

图 36-4　白毒伞内 α- 鹅膏蕈碱的化学结构

有毒的蘑菇之所以有毒，是因为在它们的体内含有毒素。毒素中最常见的是两种双环寡肽毒素：一种是 α- 鹅膏蕈碱，另外一种是毒蕈肽即鬼笔环肽（phalloidin），约 90% 蘑菇中毒致死事故都是由它们造成的。鹅膏蕈碱和毒蕈肽都是由核糖体合成的多肽前体加工而来，其中前者的多肽前体为 35 肽，后者的多肽前体为 34 肽。

α- 鹅膏蕈碱可与真核生物的 RNA 聚合酶 Ⅱ 紧密地结合，也可微弱结合 RNA 聚合酶 Ⅲ，但几乎不会与 RNA 聚合酶 Ⅰ 结合。α 鹅膏蕈碱结合在 RNA 聚合酶 Ⅱ 上一段特殊的桥螺旋的下面，横跨 RPB1 和 RPB2 这两个亚基之间的裂缝。位于桥螺旋 822 号的残基与 α- 鹅膏蕈碱上的羟脯氨酸形成氢键，此外，在 RPB1 紧靠桥螺旋的区域与毒素之间也有氢键。α- 鹅膏蕈碱并不影响 NTP 的结合，但它与聚合酶形成的这些氢键阻止了桥螺旋的移动，也就影响了 RNA 聚合酶为合成下一个磷酸二酯键必须进行的移位，使 RNA 聚合酶 Ⅱ 的移位速率降低了几千倍。

毒蕈肽的毒性比鹅膏蕈碱弱，它在细胞内能与聚合的微丝结合，但不会与肌动蛋白单体分子结合。它同聚合的微丝结合后，可抑制微丝的解体，而破坏了微丝的聚合和解聚的动态平衡，从而影响到细胞内依赖于微丝的运动系统。

若不小心误食毒蘑菇，又没有及时处理，那就十分危险！起初的反应是在食入 6~12 h 以后，肠胃不适。接着是假恢复，持续 12~24 h。由于肝细胞的功能依赖于其持续的蛋白质的合成，所以在 3~5 天以后，肝功能会受到严重伤害。严重的中毒者可能随时死亡，除非立刻进行肝移植手术。

而阻止 RNA 聚合酶的移动，因此原核生物和真核生物的 DNA 转录都可以受到它的抑制。但由于真核细胞核三种 RNA 聚合酶所催化转录的基因在 GC 含量上不同，以 rDNA 上的 GC 含量最高，因此，RNA 聚合酶Ⅰ对放线菌素 D 的抑制作用最敏感。

三、古菌的 RNA 聚合酶

古菌只有一种 RNA 聚合酶，这与细菌一样。但酶的组成和结构与真核生物的 RNA 聚合酶Ⅱ十分相似，共有 11~12 个亚基，而不像细菌只有 5 个亚基。与真核生物的 RNA 聚合酶一样，古菌的 RNA 聚合酶不受利福霉素的抑制，也不能直接识别启动子，只有在特殊的转录因子的帮助下才能识别和结合启动子。然而，古菌 RNA 聚合酶最大的亚基在 C 端没有七肽重复序列构成的 CTD。

四、某些病毒的 RNA 聚合酶

许多病毒原封不动地使用宿主细胞的 RNA 聚合酶来转录自身的基因，某些病毒则对宿主 RNA 聚合酶进行改造，以增强 RNA 聚合酶对自身基因的亲和性，而有的病毒则编码对自身基因具有高度特异性的 RNA 聚合酶，如 T7、T3 和 SP6 噬菌体，这类 RNA 聚合酶一般只由一条肽链组成。

第三节　细菌的 DNA 转录

与 DNA 复制一样，整个转录过程也可分为起始、延伸和终止三个阶段。

一、转录的起始

（一）转录起始点的确定

1. 启动子

不论是哪一种转录系统，转录都是从特定的位置开始的，即转录具有固定的起点。那么 RNA 聚合酶是如何发现正确的起点并启动基因转录的呢？通过分析比较多种基因转录起始点周围的碱基序列发现：在转录起始点的上游，存在着一些特殊的具有高度保守性的碱基序列。实验证明，在转录的起始阶段，这些保守的碱基序列作为一种标记，RNA 聚合酶能够直接或间接地识别这种标记，从而启动从特定的位点开始的基因转录，它们也因此被称为启动子（promoter）。在细菌转录系统中，RNA 聚合酶的 σ 因子能够直接识别启动子，并与之结合而启动基因的转录；但在古菌和真核转录系统之中，RNA 聚合酶并不能直接识别启动子，识别启动子的是一些特殊的转录因子。

启动子一般位于基因的上游，它与转录起始点的距离和方向都有严格的要求，不过也有一些基因的启动子位于基因的内部。

2. 确定启动子的方法

常见的确定启动子的方法有突变、序列同源性比对、电泳泳动变化分析（electrophoretic mobility shift assay，EMSA）和 DNA 酶Ⅰ－足印分析（DNase Ⅰ footprinting assay）。

EMSA 的原理是：与聚合酶特异性结合和没有结合的 DNA 在电泳时泳动的速率不同，利用这种差别可将含有启动子的和没有启动子的 DNA 片段分开，再利用足印法进行鉴定。足印法的基本原理是：与 RNA 聚合酶呈特异性结合的启动子序列可受到聚合酶的保护，而不受 DNA 酶Ⅰ的水解。如果结合 DNA 序列分析，就可以测定出受到 RNA 聚合酶保护的序列，该序列就应该属于启动子。

1975 年，David Pribnow 等人使用一种比足印法更简单的方法，测定并比较了 5 个受 RNA 聚合酶结合保护的基因片段的序列（图 36-5）。结果发现，在 -10 区有一段富含 AT 的保守序列。他们使用的方法是：先用 DNA 酶Ⅰ完全消化与 RNA 聚合酶结合的含有启动子的双链 DNA。然后，将消化过的样

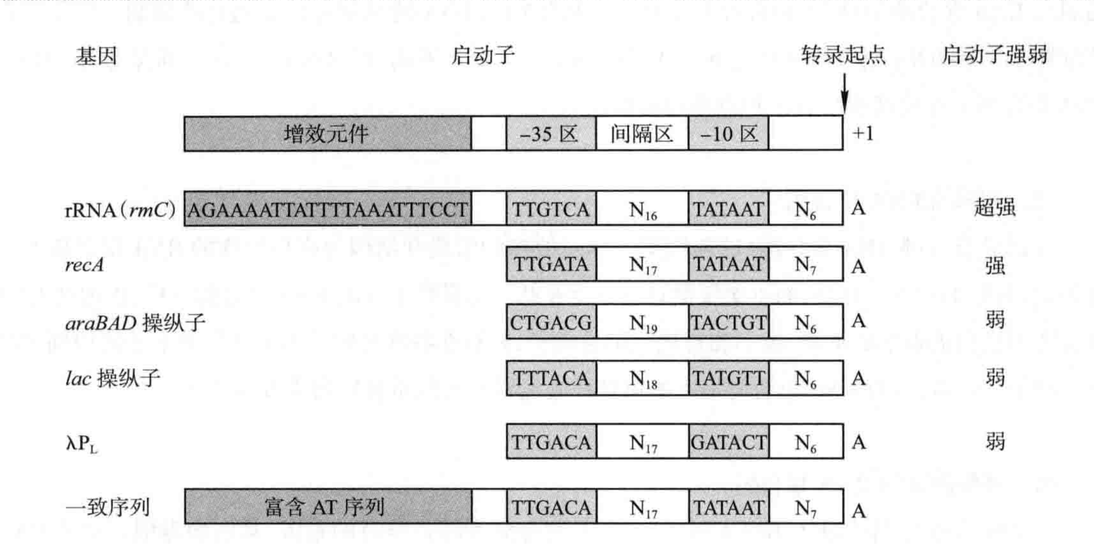


图 36-5 大肠杆菌的几种基因的启动子序列

品通过硝酸纤维素滤膜,蛋白质以及与聚合酶结合的 DNA 会与滤膜结合。随后,将吸附在滤膜上的与 RNA 聚合酶结合的 DNA 洗脱出来,再使用化学断裂法直接进行序列测定。

然而,在确定了 -10 区的保守序列以后,Pribnow 很快就发现,受到 RNA 聚合酶保护的片段在与聚合酶解离以后,无法重新与聚合酶结合,这说明单凭 -10 区尚不能保证聚合酶的结合,肯定还有其他的启动子序列。后来他们对更多位于 -10 区上游的序列进行了分析,很快发现了在 -35 区还有一段属于启动子的保守序列。

3. 细菌启动子的特征

细菌启动子的序列位于转录起始点的 5′端(图 36-6),覆盖 40 bp 左右的区域,它包含两段高度保守的序列,分别称为 -35 区和 -10 区,其中受 σ^{70} 识别的启动子在 -35 区的一致序列为 TTGACA,-10 区又名 Pribnow 盒,一致序列是 TATAAT。由其他 σ 因子识别的启动子的一致序列见图 36-6。-35 区和 -10 区之间的距离同样重要,一般为 17 ± 1 bp,原因是这样的距离可以保证这两段启动子序列处于 DNA 双螺旋的同一侧,从而有利于 RNA 聚合酶的识别和结合,否则它们会处于 DNA 双螺旋的异侧,不利于 RNA 聚合酶的识别和聚合(图 36-7)。此外,在转录活性超强的 rRNA 基因的上游 -40 区和 -60 区之间,还有一段富含 AT 的启动子序列。该序列的一致序列是 5′-AAAATTATTTT-3′,可将转录活性提高 30 倍,因此被称为增效元件(up element)。实验证明,RNA 聚合酶的 α 亚基 C 端的结构域可与该元件结合,而使 RNA 聚合酶与启动子的亲和力增强。

Quiz8 如果将一个基因的 -10 区序列突变成 ATATTA,你认为这个基因还能正常转录吗?

Quiz9 为什么 rRNA 的基因需要通过这种方式提高转录活性?

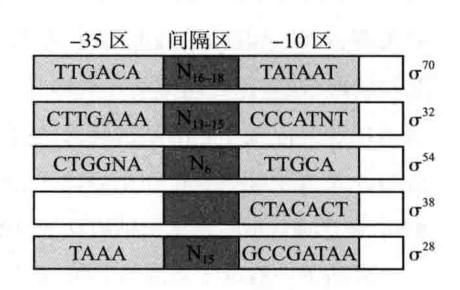


图 36-6 大肠杆菌不同的 σ 因子识别的启动子的一致序列

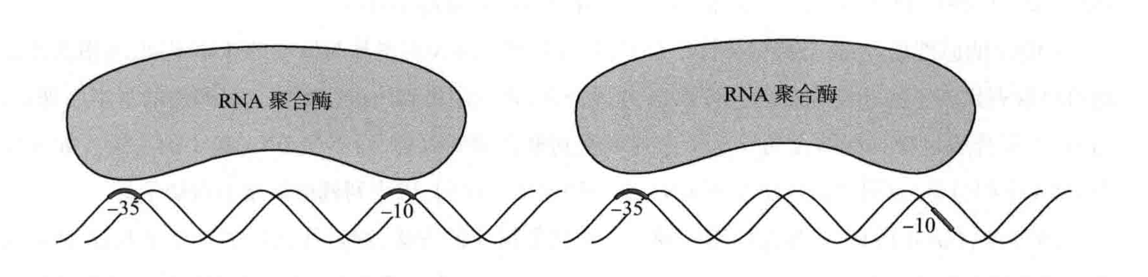


图 36-7 细菌启动子 -35 区和 -10 区之间的距离对 RNA 聚合酶识别和结合的影响

在描述启动子的位置时,碱基的位置一般以转录的起始点为参照,转录起点的位置定为 +1,位于它上游的序列为负数,下游的碱基为正数,没有 0。另外,写出的碱基序列应该属于编码链。

必须指出,启动子的一致序列是对多种基因的启动子序列的统计结果。迄今为止,在大肠杆菌中,还没有发现哪一个基因的启动子序列与一致序列完全一致。显然,一个基因的启动子序列与一致序列越相近,则该启动子的效率就越高,即属于强启动子;相反,一个基因的启动子序列与一致序列相差越大,则该启动子的效率就越低,即属于弱启动子。

有两种形式的突变可改变启动子的效率:一种是破坏与一致序列的一致性,从而导致启动子效率的降低,这样的突变称作减效突变(down mutation);另外一种是提高与一致序列的一致性,从而导致启动子效率的增强,这样的突变称为增效突变(up mutation)。

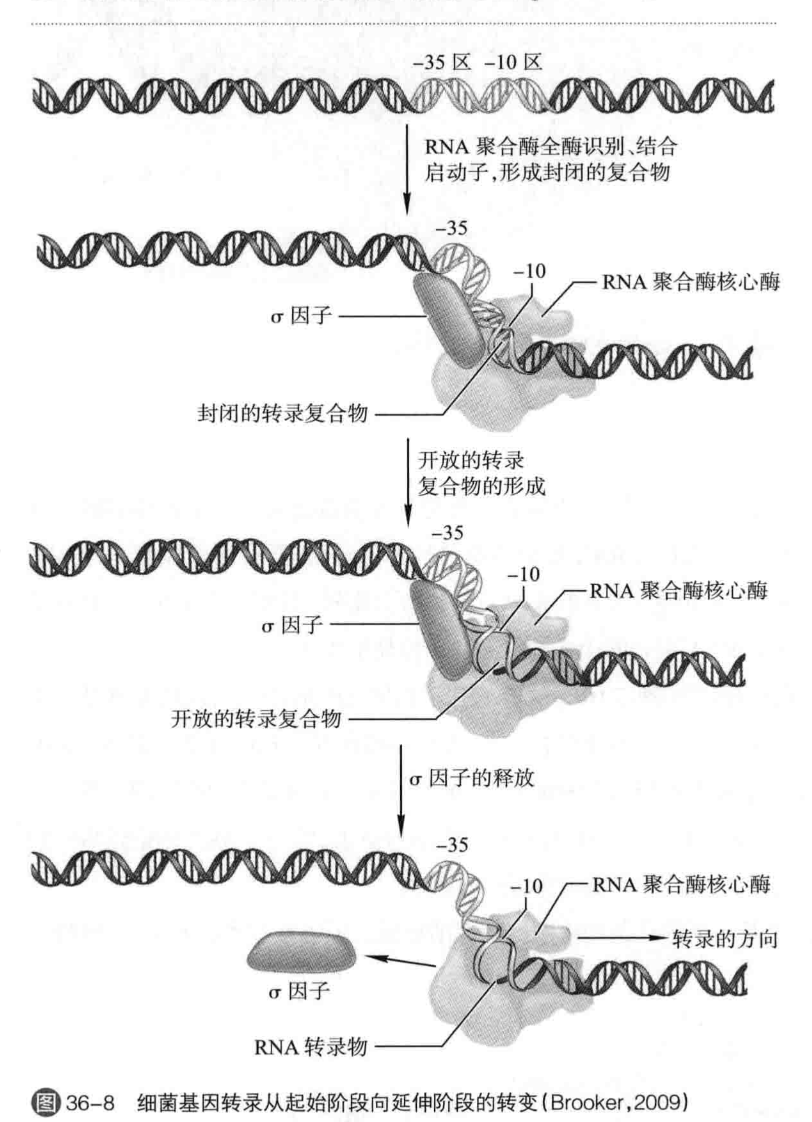


图 36-8 细菌基因转录从起始阶段向延伸阶段的转变(Brooker,2009)

(二) 转录起始复合物的形成

起始复合物的形成是转录的限速步骤,起始频率主要取决于启动子强度。强启动子每秒钟启动一次转录,弱启动子每分钟(甚至更长的时间)启动一次转录。但一旦启动成功,RNA 链延伸的速率平均为 50 nt/s,与启动子强度无关。

实际上,转录的起始是 RNA 聚合酶与启动子相互作用,并形成活性转录起始复合物的过程,整个反应可分为以下几步(图 36-8):

1. RNA 聚合酶与 dsDNA 非特异性结合

虽然 RNA 聚合酶全酶与非特异性 DNA 序列也具有一定的亲和性,但亲和性不高,因此,聚合酶只能和 DNA 随机结合。一旦结合上去,它可以在结合处沿着 DNA 向一个方向滑动、扫描,直至发现启动子序列。聚合酶与启动子的结合是特异性的,具有很高的亲和性。

2. RNA 聚合酶全酶与启动子形成封闭复合物

聚合酶在扫描中首先遇到 -35 区,与该区域的结合形成的是一种封闭复合物(closed complex)。在封闭复合物中,聚合酶主要以静电引力与 DNA 结合,DNA 也没有解链。这样的复合物并不十分稳定,半衰期为 15~20 min。足印法测定表明,在此阶段聚合酶覆盖 -55~+5 区域。

3. 封闭复合物异构化,转变成开放复合物

随着聚合酶与启动子的进一步结合,聚合酶构象发生变化,而使得 DNA 发生解链。当 DNA 解链以后,封闭复合物即转变成开放复合物(open complex),开放的复合物也就是起始转录泡(transcription bubble)。转录泡大小为 12~17 bp,一开始转录泡仅从 -10 区延伸到 -1 区,但很快以一种依赖于 Mg^{2+}

的方式从 –12 区扩展到 +2 区（图 36–9）。

开放复合物非常稳定，其半衰期在几个小时以上，这时聚合酶与启动子的相互作用不仅有静电引力，还有氢键。足印法表明在此阶段，聚合酶覆盖 –55 到 +20 区域。

开放复合物的形成是转录起始的限速步骤。现已明确，在与启动子形成复合物期间，RNA 聚合酶本身经历了显著的结构变化，σ 因子刺激封闭复合物向开放复合物的转变。开放复合物的形成不仅需要 DNA 两条链的解链，其中的 DNA 模板链还必须移至全酶的内部，以便靠近活性中心。

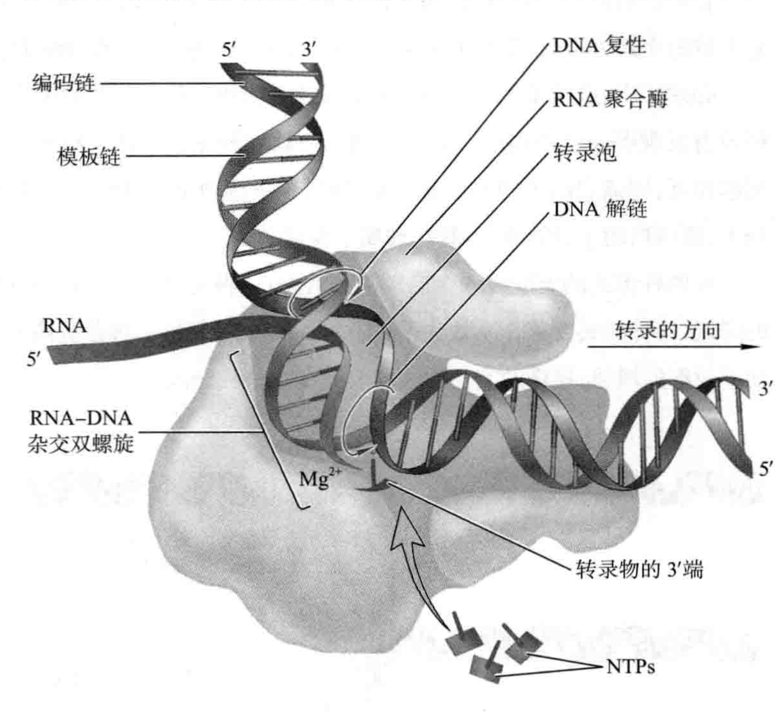


图 36–9 转录泡的结构（Brooker，2009）

4. 第一个磷酸二酯键形成

形成第一个磷酸二酯键是极为关键的一步，因为后面所有的核苷酸都是在其后依次添加的。这一步首先需要与模板链互补的前二个 NTP 与 RNA 聚合酶的活性中心结合，再由酶催化第一个 NTP 的 3′–OH 亲核进攻第二个 NTP 的 5′–α 磷酸，从而形成第一个磷酸二酯键，并释放出 PPi。一旦有了第一个磷酸二酯键，由 RNA、DNA 和 RNA 聚合酶组成的三元复合物便形成了。

Quiz10 DNA 复制和转录每参入一个核苷酸需要消耗多少个 ATP？

RNA 聚合酶催化磷酸二酯键形成的机制与 DNA 聚合酶几乎相同（图 36–10）：反应机制涉及 2 个 Mg^{2+}，一个是随 NTP 进入活性中心的，另一个本来就位于活性中心的底部，与 β′ 亚基上的 Asp460、Asp462 和 Asp464 结合。这 3 个 Asp 亚基在所有物种的 RNA 聚合酶分子上都是高度保守的。其中一个 Mg^{2+} 促进前一个核苷酸的 3′– 羟基对下一个 NTP 的 α 磷展开亲核进攻，另一个 Mg^{2+} 则促进焦磷酸基团的取代，且两个 Mg^{2+} 都有助于稳定反应中形成的磷五价过渡态。

按照上述方式，RNA 聚合酶全酶不断催化新的磷酸二酯键的形成。但通常在形成 6~10 个磷酸二

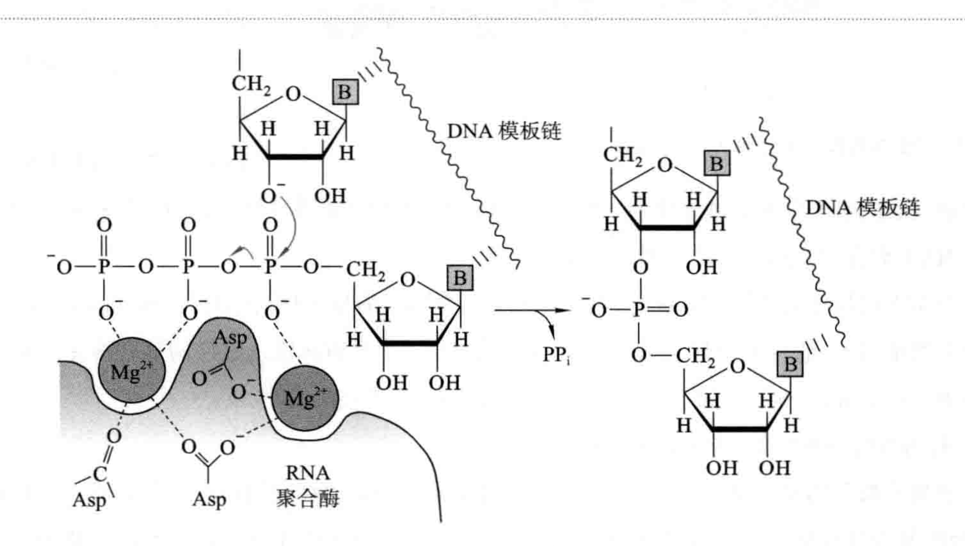


图 36–10 RNA 聚合酶催化磷酸二酯键形成的反应机制（Nelson 等，2008）

酯键以后,转录物与 σ 因子在结合核心酶的区域发生空间上的冲突,而迫使 σ 因子与核心酶解离,从此转录进入延伸阶段。解离出来的 σ 因子可以循环使用,并参与新一轮转录反应。

然而,转录要从起始过渡到延伸并非易事。在起始阶段,RNA 聚合酶会在启动子附近,重复催化短 RNA 分子的合成,这种合成称为无效转录。造成无效转录的原因是 RNA 聚合酶难以离开亲和力高的启动子,再与亲和力低的序列结合。无效转录的 RNA 长度是有限的,一般只有 3~9 nt。在发生真正意义上的延伸之前,这种重复合成可能会发生几百次。研究表明,聚合酶的活性中心能容纳 8 个碱基,这差不多等于无效转录物的长度。事实上,每合成一个无效转录物,RNA 聚合酶必须做出抉择:是离开启动子,进入延伸阶段,还是仍然结合在启动子上,继续进行无效转录。

5. 启动子清空

启动子清空(promoter clearance)是指聚合酶离开启动子,从起始过渡到延伸的过程。此过程需要解决的首要问题是,如何让聚合酶离开高亲和性的启动子,与低亲和性的非特异性 DNA 结合,从而让无效转录变成有效转录。实验证明,这是由 σ 因子控制的,正是 σ 因子赋予了聚合酶对启动子序列的特异性和高亲和性。一旦 σ 因子解离,核心酶就以一般的亲和性与 DNA 结合,转录即进入延伸阶段。

有研究表明,聚合酶在合成前 9 个核苷酸的时候,并没有在 DNA 分子上移动。只有当转录物长度达到能形成一种稳定的 RNA-DNA 杂交双螺旋的时候,RNA 聚合酶才会像尺蠖一样移动,这时启动子清空才开始进行。

二、转录的延伸

与起始阶段相比,延伸阶段的反应较为简单。失去 σ 因子的核心酶通过它的一个裂缝握住 DNA 并沿着其移动(图 36-9)。有两种相关的蛋白质——NusG 或 RfaH 结合在裂缝的两侧,形成一个完整的活动钳结构,使 RNA 聚合酶能够牢固地结合在模板上,而不会中途掉下来。这有利于维持转录的高度进行性。

在转录过程中,转录泡则随着核心酶的移动而移动,其大小维持在大约 17 bp。转录泡的维持不仅需要聚合酶在转录泡前面解链,还需要转录泡后面的 DNA 重新形成双链,转录泡内部的 RNA 链大约有 8 bp 长的片段与模板链形成 A 型双螺旋。转录泡在向前移动的时候,需要构成双螺旋的两条链的不断旋转。在旋转的过程中会不时受到一些因素的限制,例如结合在转录泡前方的各种 DNA 结合蛋白和其他结构上的障碍,使得转录泡的前后方分别形成正超螺旋和负超螺旋。正如 DNA 复制一样,不利于转录的超螺旋可被结合在转录泡前面的拓扑异构酶解除。由于转录的方向始终是从 5′→3′,新参入的核苷酸总是被添加在 RNA 链的 3′-OH 端,而每参入一个新的核苷酸,RNA-DNA 杂交双链必须发生旋转。

延伸阶段的速率并不是恒定的,已发现很多基因在转录到某些位置的时候,如富含 GC 碱基对的区域,聚合酶的移动速率趋缓,甚至发生暂停。有时,互补的 NTP 暂时短缺也可导致 RNA 聚合酶的暂停。如果这种情形发生,重新启动 RNA 合成需要内源核酸酶 GreA 和 GreB 来解除暂停状态:先是 RNA 聚合酶倒退,然后 GreA 和 GreB 切除 3′端几个核苷酸,以便让 RNA 的 3′端 -OH 能重新回到酶的活性中心。如果转录中正好发生了错误,含有错误的寡聚核苷酸也被 GreA 蛋白或 GreB 蛋白切除了,这实际上为转录提供了一种校对机制。

Quiz11 你认为在基因转录的时候,转录物在 5′端这一侧的错误更多吗?如果是的,这对一种 RNA 的功能会有影响吗?

三、转录的终止

细菌转录的终止有两种方式:一种不依赖被称为 ρ 因子(rho factor)的蛋白质,只依赖于转录物本身在 3′端自发形成的一段终止子(terminator)序列;另一种方式则依赖于 ρ 因子。

(1) 不依赖 ρ 因子的终止机制　这是细菌转录终止的主要方式,也被称为简单终止(simple termination)。该机制具有 2 个特征:一是依赖位于转录物 3′端的一串 U 序列,二是依赖紧靠 U 序列上

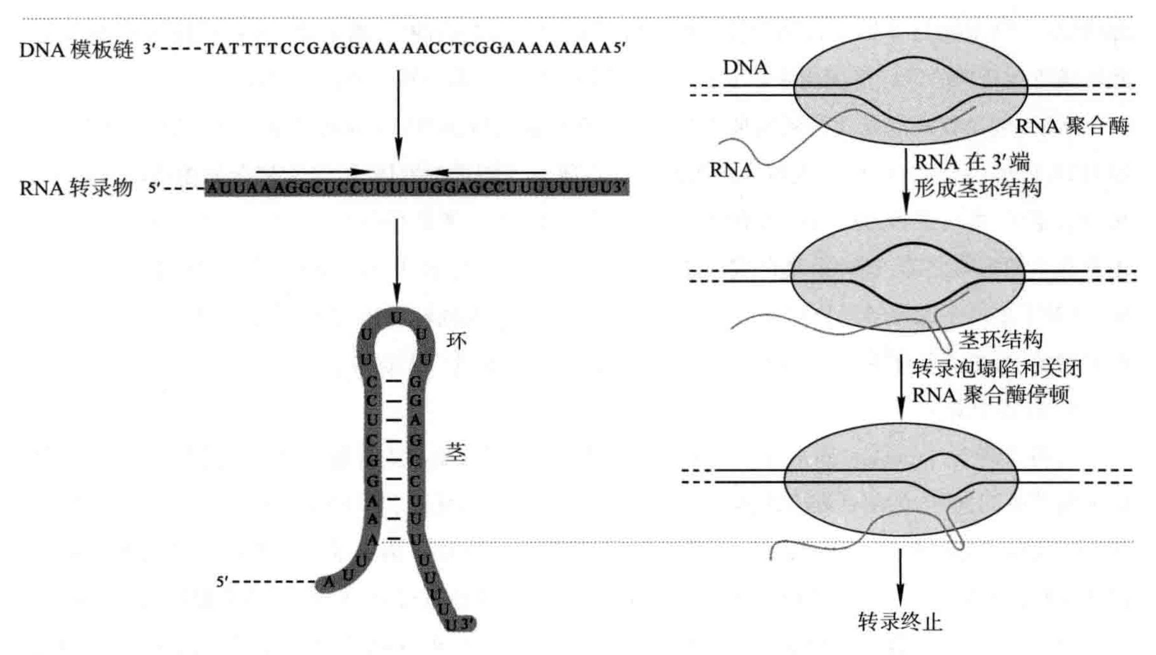


图 36-11　不依赖 ρ 因子的终止机制

游的一个富含 GC 碱基对的茎环结构(图 36-11),它们共同组成终止子的结构。

Quiz12 这种终止子结构与启动子有何相似之处?

这种终止机制的基本步骤是:当一个正在转录的 RNA 在 3′端自发形成一种富含 GC 碱基对的茎环结构后,会导致转录泡塌陷和关闭,RNA 聚合酶因此出现停顿,由于茎环结构的后面就是一串 U,与模板链以较弱的氢键结合,故很容易与模板链解离,从而导致转录物的释放和转录的终止。

(2) 依赖 ρ 因子的转录终止　这一种终止机制在细菌染色体 DNA 上较为少见,但对于噬菌体却很普遍。ρ 因子是一种同源六聚体蛋白,它具有解链酶和 ATP 酶的活性。所具有的解链酶活性使得它能够催化 RNA/DNA 和 RNA/RNA 双螺旋的解链。有证据表明,ρ 因子要发挥作用首先需要与转录物在 5′端的一段特殊的碱基序列(无二级结构,富含 C 但缺乏 G)结合,这段序列称为 ρ 因子利用位点(rho factor utilization site,*rut*)。鉴于此,有人提出了 ρ 因子作用的“热追模型”(the hot pursuit model),该模型认为(图 36-12):首先,ρ 因子识别转录物 5′端的 *rut* 位点并与之结合,以水解 ATP

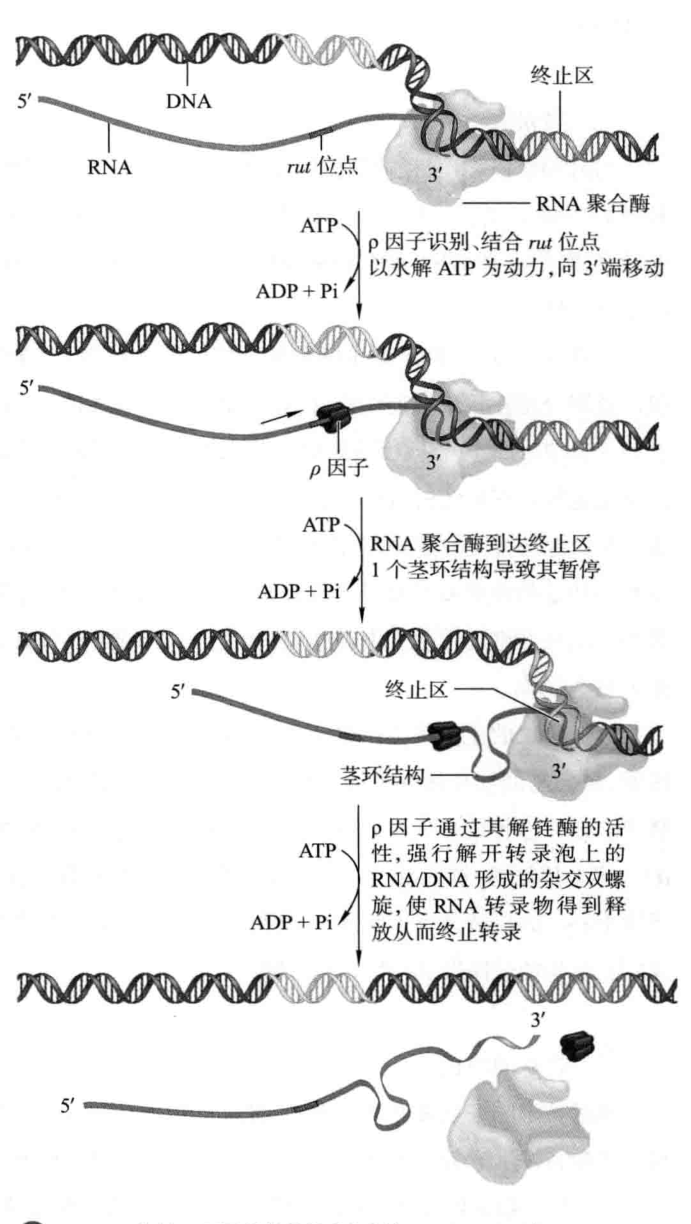


图 36-12　依赖 ρ 因子的转录终止机制(Brooker,2009)

为动力，沿着 5′→3′方向朝转录物的 3′端前进，直至遇到暂停在终止区位置的 RNA 聚合酶，暂停的原因也是由于转录物的 3′端形成茎环结构，但后面缺乏一串 U。随后，ρ 因子通过解链酶的活性，强行解开转录泡内的 RNA/DNA 杂交双螺旋，使转录物释放出来，并终止转录。

对 ρ 因子进行的 X 射线衍射分析表明，它的每一个亚基由 2 个结构域组成，其 N 端结构域结合 RNA，C 端结构域结合 ATP，在结构上和参与氧化磷酸化的 F_1F_0–ATP 合酶的 α 和 β 亚基同源。整个 ρ 因子的三维结构呈锁紧垫圈状的螺旋。螺旋沿着螺旋轴约上升 4.5 nm，直径为 12 nm，内部有一个直径约为 3 nm 的孔。首尾两个亚基在没有结合 RNA 的时候有 1.2 nm 的空隙，但一旦结合 RNA，空隙立刻消失，这时候的 ρ 因子就像一个密闭的环套在 RNA 链上。每 2 个亚基构成 1 个单位，因此 1 个 ρ 因子共有 3 个单位。ρ 因子在作用的时候，3 个单位在某一个瞬间的构象并不相同，而是通过结合和水解 ATP 使构象交替发生改变，这类似于 F_1F_0–ATP 合酶在合成 ATP 时，所使用的“结合变化”机制，于是 ρ 因子像一个分子马达一样，沿着 mRNA 轨道朝着 3′端进发。

某些蛋白质能够干扰 ρ 因子的终止作用，并以此作为调节基因表达的一种手段，如 λ 噬菌体表达的 N 蛋白就有这样的功能（参看第四十二章“原核基因表达的调控”）。

第四节　真核生物核基因的转录

一、真核细胞核转录系统与细菌转录系统的异同

虽然这两大转录系统在转录的基本机制上十分相似，但真核生物具有以下几点特殊的性质：

1. 染色质和核小体结构对转录有深刻的影响

真核生物的核基因处于核小体和染色质结构之中。可以预见，某个基因在转录之前或转录之中，其所在位置的核小体结构和染色质构象必须发生某种有利于转录的变化，即染色质发生重塑(remodeling)，使染色质构象从紧密状态变为松散状态。因为只有这样，参与转录的一系列酶和蛋白质才能有效地识别启动子和模板等，并形成有功能的转录复合物（参看第四十三章“真核生物的基因表达调控”/ 第一节“在染色质水平上的基因表达调控”）。

2. RNA 聚合酶高度分工（参见本章“RNA 聚合酶的结构与功能”）

3. 转录还需要许多被称为转录因子的蛋白质的参与

转录因子有基础转录因子（basal transcription factor）和特异性转录因子（specific transcription factor），前者是所有的基因转录所必需的，也称为普遍性转录因子（general transcription factor），而后者为特定的基因转录所必需的，也称为调控性转录因子（regulatory transcription factor）。对于 RNA 聚合酶Ⅰ和Ⅲ负责转录的基因，在所有的细胞中都表达，因此所需要的转录因子一般都是基础转录因子。而 RNA 聚合酶Ⅱ催化转录的基因，有的在所有的细胞都表达，有的则具有组织特异性。组织特异性表达的基因显然需要特异性转录因子。

基础转录因子的功能包括：作为装配因子或定向因子去识别和结合核心启动子，招募 RNA 聚合酶Ⅱ正确地与启动子结合，促进 DNA 的解链及调节 RNA 聚合酶的活性；与其他上游元件或反式作用因子结合或相互作用；参与蛋白质与蛋白质的相互作用，有助于转录起始复合物的装配和稳定。

4. 启动子以外的序列参与调节基因的转录

这些特殊的序列因为与转录的基因位于同一条染色体 DNA 上，呈顺式关系，所以一般通称为顺式作用元件（*cis*-acting element）。然而，顺式作用元件单独起不了作用，只有在结合了特殊的蛋白质因子以后才行。就结合在顺式作用元件上的蛋白质因子而言，它们本身的基因一般位于其他的染色体 DNA 分子上，与调节的基因呈反式关系，故被称为反式作用因子（*trans*-acting factors）。

5. 转录与翻译不存在偶联关系

细菌的 mRNA 一旦开始合成，核糖体就与其 5′端结合，启动翻译，在转录与翻译之间存在偶联关系。但真核细胞的转录和翻译分别发生在细胞核和细胞质，在时空上完全分割，没有偶联关系。

6. 转录的产物多为单顺反子，而原核基因的转录产物多为多顺反子

细菌功能相关的基因通常共享一个启动子，以一个共同的转录单位进行转录，因此产物为多顺反子。而真核生物的绝大多数基因都有自己的启动子，独立地进行转录，因此产物为单顺反子。

二、真核细胞核 DNA 的转录

真核细胞核内的 3 种聚合酶负责不同性质基因的转录，现分别加以讨论。

（一）RNA 聚合酶 I 所负责的基因转录

RNA 聚合酶 I 在核仁催化 28S rRNA、18S rRNA 和 5.8S rRNA 的转录。这 3 种 rRNA 的合成受同一个启动子控制，45S rRNA 是它们的共同前体。

1. 启动子

这类启动子本来就属于强启动子，但具有一定的物种特异性（species specificity），即某一物种的启动子只对本物种的基因转录有效，对其他物种无效。在结构上，它包括两个部分：其一为位于 −31 区到 +6 区之间的核心启动子（core promoter）——与基因的基础转录有关；其二是位于 −187 区到 −107 区之间的上游启动子元件（upstream control element，UCE）——与基因的有效转录有关。

UCE 与核心启动子的序列高度同源，约有 85% 的序列相同。两者都富含 GC（60%），但核心启动子在转录起始点附近却倾向于富含 AT，这显然有利于 DNA 双螺旋的解链。

2. 基础转录因子——TFI

RNA 聚合酶 I 需要 2 种转录因子：一种是 UCE 结合因子（UCE binding factor，UBF），它为单一的多肽，无物种特异性，负责识别、结合 UCE 和核心启动子上富含 GC 的序列；另一种是选择因子 1（selectivity factor1，SL1），具有物种特异性，由多个亚基组成，包括 1 个 TATA 盒结合蛋白（TATA box binding protein，TBP）和 3 个 TBP 相关因子（TBP associated factor，TAF），其中 TBP 为 3 种 RNA 聚合酶所共享。

3. 转录的起始和延伸

如图 36-13 所示，在转录的起始阶段，UBF 作为组装因子（assembly factor）与核心启动子和 UCE 结合。一旦 UBF 结合到启动子上，SL1 随即被招募上去。SL1 的结合一方面稳固了 UBF 和启动子的结合，另一方面又作为定位因子（positional factor），引导聚合酶 I 正确地定位到启动子上，为转录从正确的起始点开始创造了条件。没有 UBF，rDNA 也能转录，但效率会大大降低。

4. 转录的终止

这类基因的转录终止于一个由 18 nt 组成的分散的终止子区域，该终止子序列位于编码区末端序列下游约 1 000 nt 处，需要一种 DNA 结合蛋白（小鼠为 TTF-1）充当转录终止因子，具体机制尚不十分

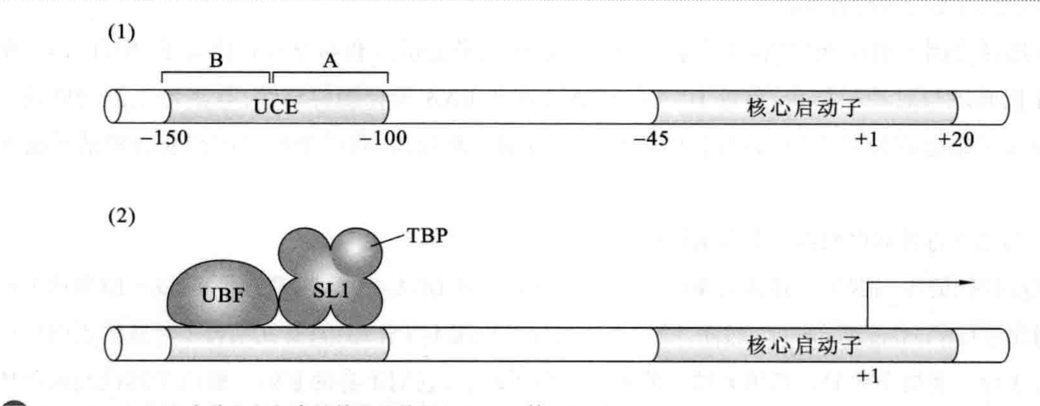


图 36-13　RNA 聚合酶 I 所负责的基因的转录（Watson 等，2007）

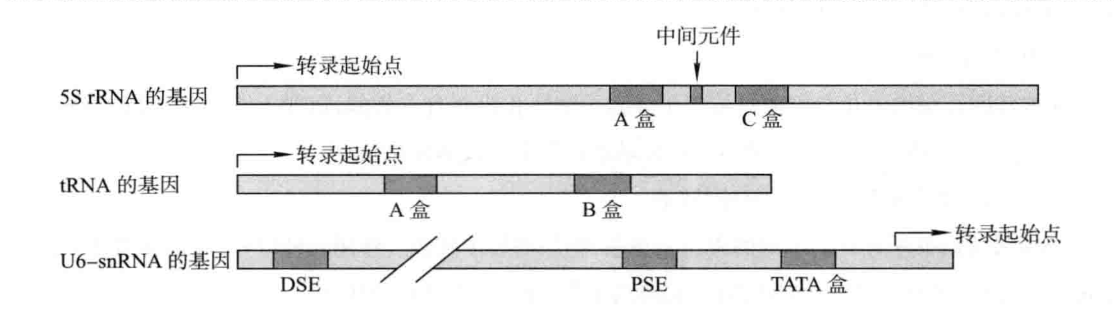


图 36-14 由 RNA 聚合酶Ⅲ催化转录的基因的启动子结构

清楚。在小鼠细胞内，一旦 RNA 聚合酶 I 遇到结合在终止子上的 TTF-1，终止便发生了：先是 TTF-1 招募 1 种释放因子，催化 3′端的形成；然后，1 种外切酶可剪切 rRNA 前体新生的 3′端，产生成熟的 3′端；最后，RNA 聚合酶 I 与模板解离。

（二）RNA 聚合酶Ⅲ所负责的基因转录

该酶负责稳定的小 RNA 的转录，如 tRNA、5S rRNA、7S RNA 和无帽子结构的 snRNA。

1. 启动子

此类启动子分为两种类型（图 36-14）：一类与 RNA 聚合酶Ⅱ相似，也位于基因的上游，含有 TATA 盒、近端序列元件（proximal sequence element，PSE）和远端序列元件（distal sequence element，DSE），如 7SK RNA、7SL RNA 和 U6 snRNA；另一类位于基因内部，属于内部启动子（internal promoter），如 tRNA、5SL rRNA 和腺病毒的 VA RNA。绝大多数 tRNA 基因的启动子包括 A 盒（A Box）和 B 盒（B box），两盒分别对应于 tRNA 的 D 环和 TψC 环，但 $tRNA^{Sec}$ 的基因也含有 DSE。A 盒和 B 盒的保守序列分别是 5′-TGGCNNAGTGG-3′和 5′-GGTTCGANNCC-3′。由于它们是基因内启动子，因此本身也被转录。5S rRNA 的启动子包括 A 盒、中间元件（intermediate element，IE）和 C 盒。

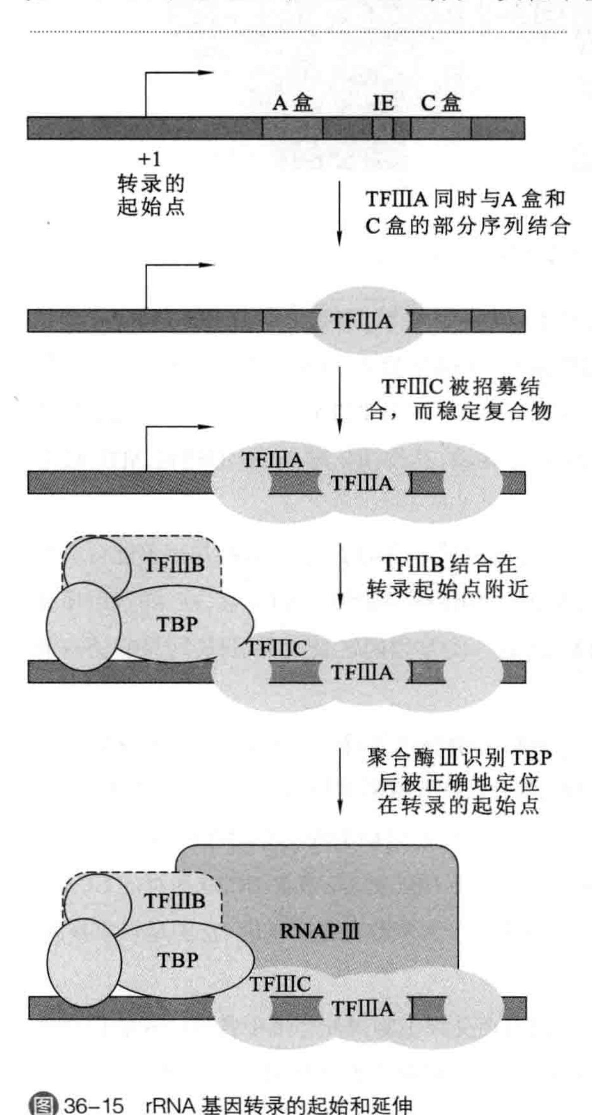


图 36-15 rRNA 基因转录的起始和延伸

2. 基础转录因子——TFⅢ

TFⅢ有 3 种，即 TFⅢA、B 和 C。其中 TFⅢC 为组装因子，负责与 A 盒和 B 盒结合。TFⅢB 是一种定位因子，结合在 A 盒上游约 50 bp 的位置，但与它结合的序列无特异性，这说明 TFⅢB 结合的位置是由 TFⅢC 决定的。TFⅢB 由 TBP、TFⅢB 相关因子（TFⅢB-related factor，BRF）和 TFⅢB″ 组成。TFⅢA 由一条肽链组成，有锌指结构，它仅为 5S rRNA 基因的转录所必需的组装因子。

3. 转录的起始和延伸

以 5S rRNA 基因的转录为例（图 36-15），首先是 TFⅢA 与启动子结合；然后，TFⅢA 将 TFⅢC 招募进来，形成一种稳定的复合物，TFⅢC 几乎覆盖整个基因；随后，TFⅢC 再将 TFⅡB 招募到转录起始点附近；最后，TBP 将 RNA 聚合酶Ⅲ招募进

来，形成有功能的转录起始复合物。

4. 转录的终止

这类基因转录的终止与细菌不依赖于 ρ 因子的终止机制有点相似，需要 1 段富含 GC 的序列和 1 小串 U，但 U 的长度短于细菌，且富含 GC 区域也不需要形成茎环结构。

（三）RNA 聚合酶Ⅱ所负责的基因转录

RNA 聚合酶Ⅱ所催化的基因的表达最复杂，除了需要多种顺式作用元件以外，还需要多种不同的反式作用因子，其中后者包括激活蛋白、辅激活蛋白、介导蛋白和转录因子。

1. 启动子与其他控制基因转录的顺式元件

与 mRNA 的基因转录有关的顺式作用元件包括：核心启动子、调控元件（regulatory element）、增强子和沉默子。

（1）核心启动子　又名基础启动子（basal promoter），其功能与原核生物的启动子相当，负责招募和定位 RNA 聚合酶Ⅱ到转录起始点，以正确地启动基因的转录，也可通过促进转录复合物的装配或稳定转录因子的结合而提高转录的效率。

属于核心启动子的有（图 36–16）：TATA 盒、起始子（initiator，*Inr*）、TFⅡB 识别元件（TFⅡB recognition element，BRE）、模体十元件（motif ten element，MTE）和下游启动子元件（downstream promoter element，DPE）。

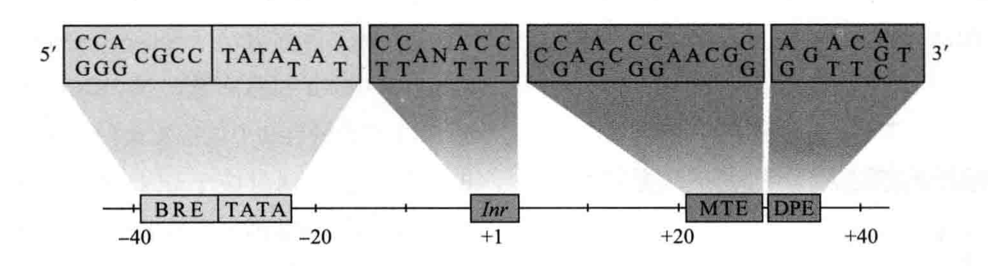


图 36–16　RNA 聚合酶Ⅱ催化转录的基因可能所具有的核心启动子序列（Voet 等，2011）

TATA 盒为一段富含 AT 碱基对的核苷酸序列，位于 –30~–25 区域，与细菌启动子的 Pribnow 盒相似，但位置不同；*Inr* 覆盖转录的起始点，通常 –1 碱基为 C，+1 碱基为 A。TATA 盒和 *Inr* 属于招募和定位元件，主要用来决定转录的起点；BRE 位于 –37~–32 区域，可视为 TATA 盒向上游的扩展，是转录因子 TFⅡB 的识别序列；DPE 位于 *Inr* 下游，总在 +28~+32 区域，其作用依赖于 *Inr* 的存在；MTE 位于 +18~+27 区域，它的作用也依赖于 *Inr* 的存在。

需要特别强调的是，并非所有蛋白质的基因都有上述 5 种核心启动子元件，有些基因可能含有其中的某些元件，有些基因也可能都缺乏。据估计，大概 2/3 的蛋白质基因无 TATA 盒，1/2 的蛋白质基因无 *Inr*。那些既无 TATA 盒、又无 *Inr* 的基因数目甚少，这一类基因的转录效率很低，转录起始点的位置也不固定。

（2）调控元件　其作用是调节基因的转录效率，包括上游临近元件（upstream proximal element，UPE）和上游诱导元件（upstream inducible element，UIE），两者的作用都需要结合特殊的反式作用因子。

UPE 一般为一些短的核苷酸序列，长度在 6~20 nt，因位于核心启动子的上游附近而得名，它的存在能影响转录起始的效率，但不影响转录起点的特异性。属于 UPE 的有：富含 GC 碱基对的 GC 盒、CCAAT 盒、AP2 盒和八聚核苷酸元件等。其中 GC 盒通常位于大多数蛋白质基因（特别是管家基因）的上游，而且往往有多个拷贝。

UIE 存在于核心启动子的上游。含有 UIE 的基因只有在受到细胞内外环境中各种特殊的信号诱导才表达，或者表达增强。常见的 UIE 有：激素应答元件 HRE、热激应答元件 HSE、cAMP 应答元件 CRE、金属应答元件 MRE 和血清应答元件 SRE 等。

(3) 增强子和沉默子　增强子(enhancer)是一种能够大幅度增强基因转录效率的顺式作用元件，而沉默子(silencer)正好与增强子相反，是一种阻止基因表达的顺式作用元件。它们的作用都需要特定反式作用因子的结合，其中与增强子结合的反式作用因子叫激活蛋白，与沉默子结合的叫阻遏蛋白。

与启动子不同的是，增强子和沉默子的作用具有以下特性：①与距离无关——既可以在距离基因很近的地方，也可以在很远的地方发挥作用；②与方向无关——相对于基因的方向可随意改变而不影响其作用效率；③与位置无关——既可以在基因的上游也可以在基因的下游，甚至可以在基因的内部发挥作用；④对临近的基因作用最强；⑤某些增强子或沉默子具有组织特异性。

为了解释为什么增强子或沉默子能够在距离启动子很远的地方起作用，有人提出了环出模型(looping-out model)。该模型认为，增强子或沉默子在作用的时候，位于启动子与增强子或启动子与沉默子之间的碱基序列会以环的形式突出来。如果是增强子，那么与增强子结合的激活蛋白就通过辅激活蛋白(co-activator)，与基础转录因子和 RNA 聚合酶相互作用，促进 RNA 聚合酶和转录因子与启动子的结合，从而促进基因的转录(图 36-17)；如果是沉默子，那么与沉默子结合的阻遏蛋白就通过辅阻遏蛋白(co-repressor)，与基础转录因子和 RNA 聚合酶相互作用，抑制基因的转录。

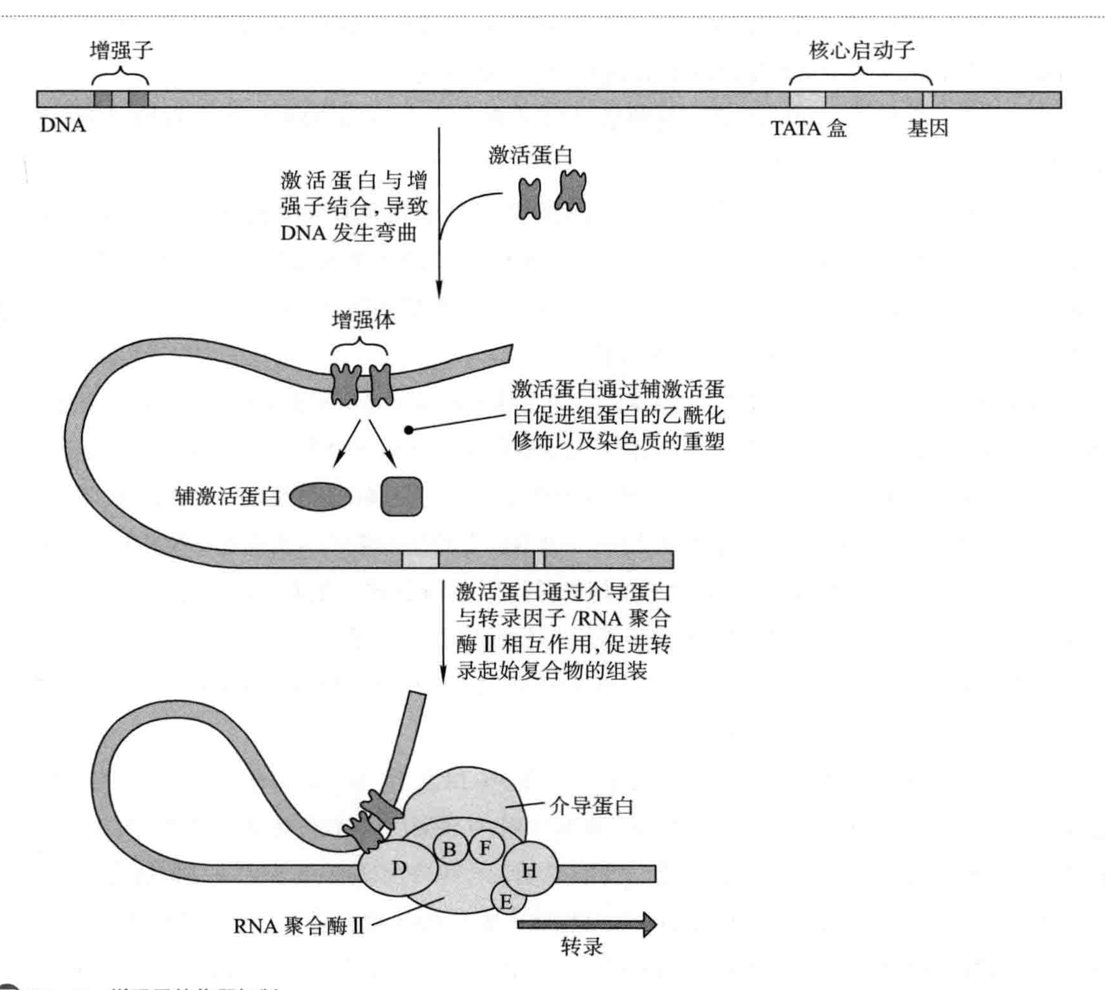


图 36-17　增强子的作用机制

在一个基因的表达被激活的时候，一系列的激活蛋白和辅激活蛋白被招募到它的增强子上，形成一种高度有序的增强子与蛋白质的三维复合物，此复合物称为增强体(enhanceosome)。增强体的组装具有协同性，它通过激活蛋白与增强子的结合，将组蛋白修饰酶、染色质重塑因子和基础转录因子有序地招募到启动子的周围，从而激活基因的表达。

2. 基础转录因子——TFⅡ

参与蛋白质基因转录的转录因子有基础转录因子和特异性转录因子。前者为所有的蛋白质基因

转录所必需，后者为特定的蛋白质基因转录所必需。

RNA 聚合酶Ⅱ所需要的基础转录因子有 TFⅡA、B、D、E、F、H、S 和 J 等。

Quiz13 为什么 TBP 的突变通常是致死性的？

(1) TFⅡD 是一种多亚基蛋白，由 TBP 和 TAF 组成。TBP 有 2 个非常相似的马镫状结构域。这两个马镫状结构域由一个短的碱性肽段相连，这段碱性肽段在小沟与 DNA 结合，两侧的马鞍状结构域横跨在 DNA 分子上。TBP 与小沟的结合可导致 DNA 出现 80°的弯曲。TFⅡD 的功能包括：识别和结合核心启动子；结合和招募其他基础转录因子；作为多种调节蛋白的作用目标；具有激酶（磷酸化 TFⅡF）、组蛋白乙酰转移酶和泛素激活酶 / 结合酶活性。

(2) TFⅡA 由 3 个亚基组成，其功能包括：与 TBP N 端的马镫状结构域结合；取代与 TBP 结合的负调控因子（如 NC1 和 NC2/DR1）；稳定 TBP 与 TATA 盒的结合。

(3) TFⅡB 只由一条肽链组成，其功能包括：与 TBP C 端的马镫状结构域结合；与 TFⅡF 的 RAP30 亚基结合，从而将聚合酶Ⅱ和 TFⅡF 形成的复合物招募到启动子；稳定 TBP 与 DNA 的结合；作为多种激活蛋白的作用目标。

(4) TFⅡF 由 RAP38 和 RAP74 亚基组成，其功能包括：与 TFⅡB 和聚合酶Ⅱ结合，降低聚合酶与非特异性 DNA 的相互作用，促进聚合酶与启动子的结合；参与转录的起始和延伸，降低延伸过程中的暂停，保护延伸复合物免受阻滞；RAP74 具有解链酶活性，可能参与启动子的解链以暴露模板链。在无 TFⅡB 的情况下，可刺激磷酸酶的活性，促进 CTD 的去磷酸化。

(5) TFⅡE 由 2 个亚基组成，其功能包括：与聚合酶Ⅱ结合，招募 TFⅡH；调节 TFⅡH 的解链酶、ATP 酶和激酶活性；参与启动子的解链。

(6) TFⅡH 有 9 个亚基，其中有一个亚基是周期蛋白 H，同时具有 ATP 酶、解链酶和蛋白质激酶的活性。其功能包括：与 TFⅡE 紧密相连，相互结合和调节；其 2 个最大的亚基（XPB 和 XPD）所具有的解链酶活性促进转录过程中模板的解链；为第一个磷酸二酯键的形成所必需；激酶的活性导致聚合酶Ⅱ的 CTD 的磷酸化修饰，从而促进启动子的清空。

Quiz14 TFⅡH 除了参与 DNA 转录，还参与哪一种过程？

(7) 转录延伸因子 经纯化过的 RNA 聚合酶Ⅱ在体外转录的速率只有 100~300 nt/min，远远低于体内转录的实际速率（1 200~2 000 nt/min），而且转录断断续续，经常发生暂停。之所以如此，是因为体内的转录延伸有转录延伸因子的帮助。有两类延伸因子：一类是基础延伸因子——能加快所有蛋白质基因的有效转录；另一类是调节性延伸因子——调节特定的基因转录，人们对这一类还知之甚少。

转录延伸因子所起的作用是防止转录暂停或提前终止，促进转录在染色质上的延伸。它们对于比较大的基因的转录和某些在发育早期需要快速表达的基因尤为重要。

一些重要的基础延伸因子也已被鉴定，它们包括 P-TEFb、TFⅡS、延伸蛋白（elongin）和 ELL。在这些延伸因子之中，TFⅡS 被研究得最为详尽。

TFⅡS 的功能相当于大肠杆菌的 GreB 蛋白。如图 36-18 所示，与细菌一样，真核生物在转录延伸过程中，RNA 聚合酶Ⅱ有可能遇到一些特别的碱基序列，导致酶后退，致使转录物的 3′端突出，进而使转录暂停。暂停的解除需要 TFⅡS 的帮助。TFⅡS 含有两个保守的区域：一个中央结构域为结合 RNA 聚合酶Ⅱ所必需；另一个含有锌指模体的 C 端结构域可刺激 RNA 聚合酶Ⅱ切下在 3′端突出的几个核苷酸。X 射线衍射研究表明，其锌指结构上突出一个 β 发夹结构，正好与聚合酶活性中心互补，其中的两个不变、为发挥功能所必需的酸性氨基酸与聚合酶Ⅱ的活性中心一起装配成外切酶活性中心，而与发夹结构结合的金属离子可能极化水分子，有利于对磷原子的亲核进攻。RNA 聚合酶的外切酶活性不但可以解除转录的暂停，还可能用来切除错误参入的核苷酸，进行转录校对。

3. 激活蛋白与辅激活蛋白

激活蛋白与直接与增强子结合，增强基因的表达效率，而辅激活蛋白也称为辅激活物，是辅助激活蛋白起作用的一大类蛋白质的总称，主要包括组蛋白乙酰转移酶复合物和促进染色质构象松散的各种染色质重塑因子。

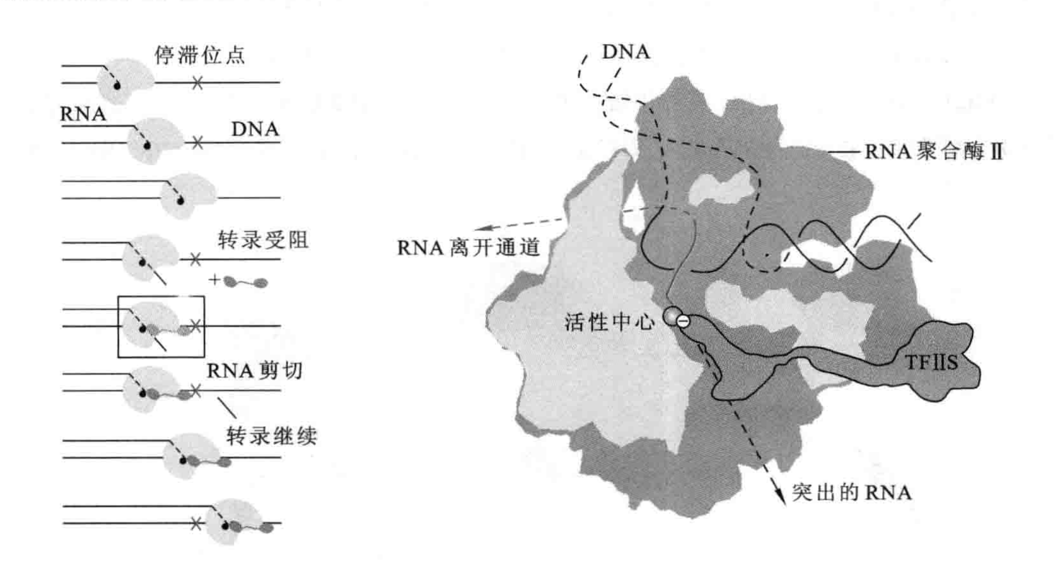


图 36-18 TFⅡS 诱导的 mRNA 剪切

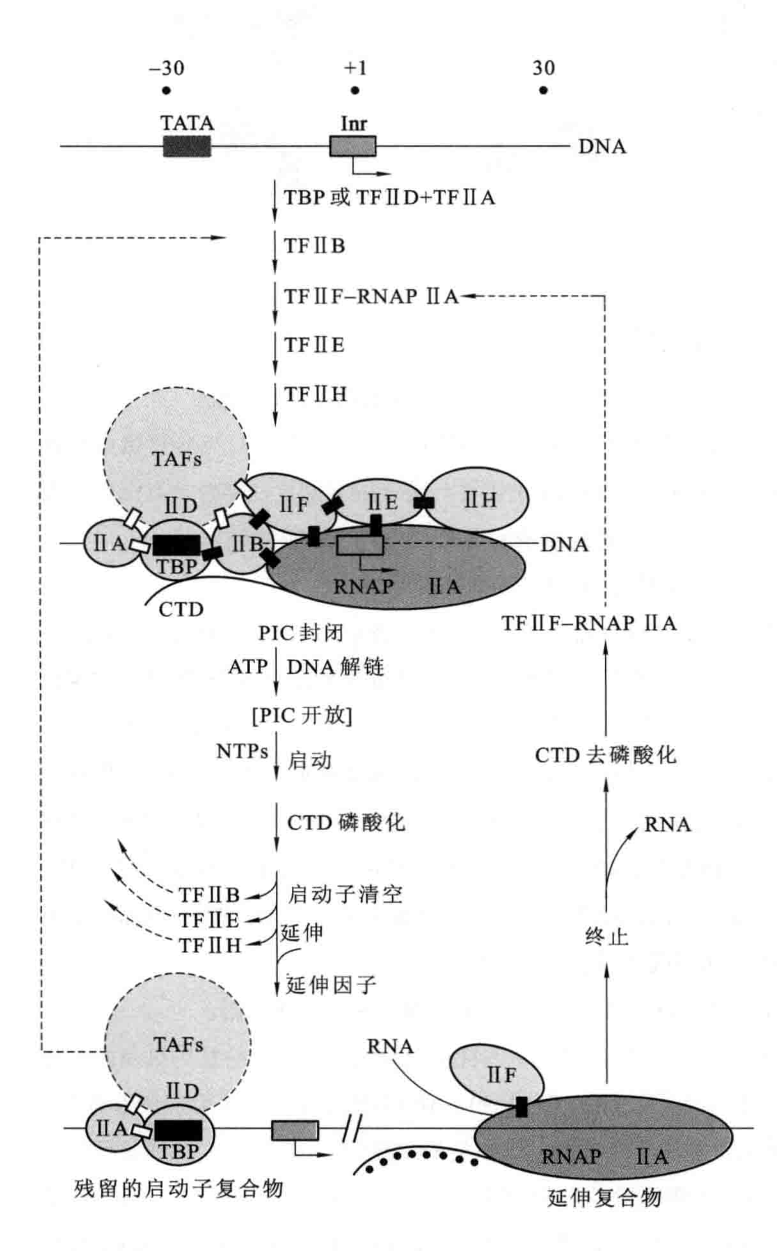


图 36-19 RNA 聚合酶Ⅱ催化的基础转录的全过程(Mathews 等,1999)

4. 介导蛋白

介导蛋白(mediator)是在纯化 RNA 聚合酶Ⅱ时得到的一类与 CTD 结合的庞大的蛋白质复合物,约由 20 种蛋白质组成,有时被归入辅激活蛋白一类。它们在细胞中的功能是充当激活蛋白和阻遏蛋白与 RNA 聚合酶/基础转录因子相互作用的桥梁或纽带。无论是激活蛋白还是阻遏蛋白,都需要通过介导蛋白才能将激活或阻遏基因转录的信息传递给 RNA 聚合酶/基础转录因子。

介导蛋白有 2 类:一类为 SRB 蛋白,它们直接与 CTD 结合,可校正 CTD 的突变;另一类为 SWI/SNF 蛋白,其功能是破坏核小体结构,促进染色质的重塑。

5. 转录的起始、延伸和终止

一个蛋白质基因的转录可以在没有激活蛋白和辅激活蛋白的存在下进行,但这种情况下的转录效率将非常低,这种最低水平的转录一般被称为基础转录。基础转录的起始是由各基础转录因子和 RNA 聚合酶Ⅱ按照一定的次序(图 36-19),通过招募的方式形成 PIC。转录因子和聚合酶Ⅱ

与启动子结合的大致次序是：TFⅡD → TFⅡA → TFⅡB →（TFⅡF + RNAP Ⅱ）→ TFⅡE → TFⅡH。

以含有 TATA 盒和 *Inr* 的启动子为例，一轮转录循环包括以下几步反应：

（1）PIC 的形成　先是 TFⅡD 识别并结合 TATA 盒，随后是 TFⅡA 和 TFⅡB 被招募进来，然后是 TFⅡF/ 聚合酶Ⅱ和介导蛋白被招募进来，最后是 TFⅡE 和 TFⅡH 的结合，形成完整的 PIC（图 36–20）。

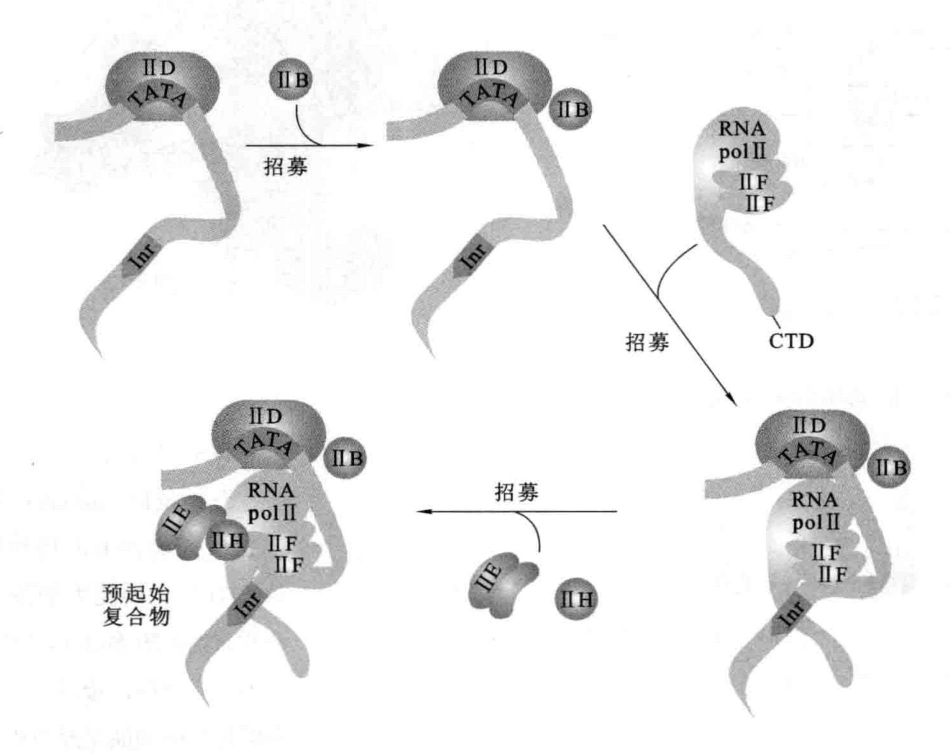


图 36–20　RNA 聚合酶Ⅱ催化的基因转录预起始复合物的形成

（2）PIC 从封闭状态转变成开放状态　在开放状态下，DNA 已解链，RNA 开始合成。

（3）CTD 的磷酸化　催化 CTD 磷酸化的蛋白质激酶有：TFⅡH 分子中的 CDK7，与介导蛋白相连的 CDK8，与转录因子 P–TEFb 相连的 CDK9。这些 CDK 都需要与特定的周期蛋白结合才有活性。在 CTD 高度磷酸化以后，介导蛋白与 CTD 解离。同时高度磷酸化的 CTD 与 TBP“脱钩”。

（4）RNA 聚合酶Ⅱ离开启动子，启动子被清空，转录进入延伸阶段。

启动子清空以后，TFⅡE、TFⅡH、TFⅡA、TFⅡD 和介导蛋白仍然留在启动子上，TFⅡS 则和 TFⅡF 一起继续参与延伸。图 36–21 显示了由 RNA 聚合酶Ⅱ催化的转录延伸复合物的三维模式结构，按照惯例，聚合酶沿着 DNA 模板从 3′→ 5′端移动。进入酶的 DNA 被蛋白质钳子或颌（jaw）握住，受舵（rudder）的作用发生解链，RNA 的 3′端紧靠活性中心的 Mg^{2+}。蛋白质壁（wall）阻断核酸笔直穿过酶，结果导致 DNA–RNA 杂交双链的轴与进入的 DNA 之间近乎垂直。如此弯曲将 DNA–RNA 杂交双链的末端暴露给从漏斗状孔进入的 NTPs，同时迫使 RNA 在 5′端与 DNA 分开，且只能从舵和蛋白质环状盖子（lid）下方的通道离开，防止 DNA–RNA 杂交双链的长度超过 9 bp。舵和盖子从一个大的钳子（clamp）中出来，钳子跨过活性中心区域，夹住 DNA，从而提高转录的进行性。

（5）当转录到一定阶段，受到 TFⅡF 的刺激，CTD 上的磷酸基团被一种磷酸酶水解。

（6）转录终止　转录终止于一段终止子区域，但终止子的性质以及它如何影响终止的机制尚不知晓。已有证据表明，与 CTD 结合的参与加尾反应的 CPSF 和 CstF 可能在调节终止反应中起作用。原因是它们在加尾反应中的作用是先识别位于转录物 3′端附近的加尾信号，然后再在加尾信号的下游某一位置引入切口，为多聚 A 聚合酶的作用做准备（参看第三十七章“转录后加工”）。显然，转录物在 3′端受到剪切以后，转录是很难继续下去了，自然地进入终止阶段。而转录一旦终止，RNA 聚合酶Ⅱ/TFⅡF 复合物则离开模板，与介导蛋白再结合，重新启动下一轮的转录。

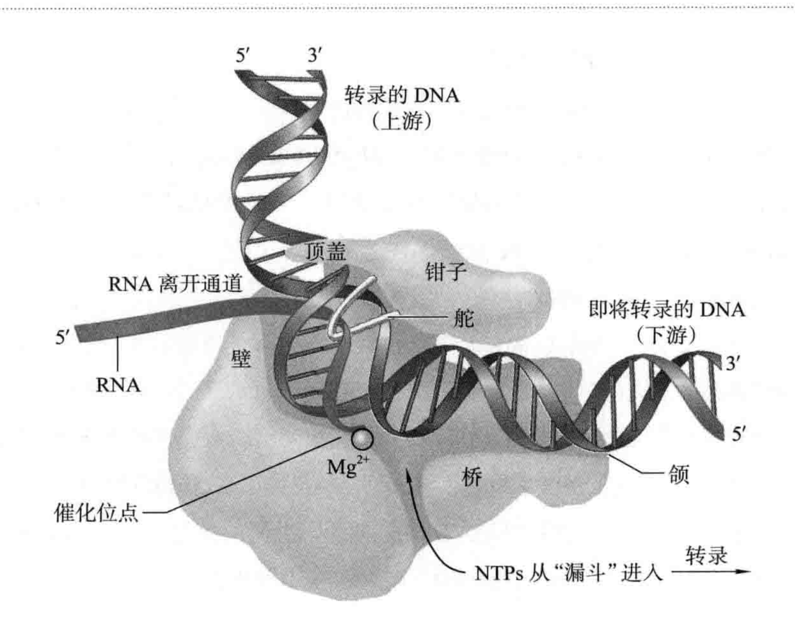


图 36-21 RNA 聚合酶Ⅱ催化的转录延伸

第五节 古菌的 DNA 转录

与复制系统相比，古菌似乎拥有简版的真核转录系统。无论是催化转录的 RNA 聚合酶的结构和性质，还是启动子的结构以及转录的基本过程都与真核生物极为相似。然而，在基因表达的调控方面，古菌与细菌接近（参看第四十二章"原核生物的基因表达调控"）。

在启动子的结构上，古菌类似于真核生物中由 RNA 聚合酶Ⅱ所负责转录的基因，一般由 3 个部分组成：一是位于转录起点上游的 TATA 盒，它由 TBP 识别；二是位于 TATA 盒上游的 BRE，它由 TFB 识别；三是位于转录起点的起始子元件（*Inr*）。

古菌 DNA 转录起始的基本过程是（图 36-22）：首先是 TBP 结合 TATA 盒，然后是 TFB 结合 BRE，随后 RNA 聚合酶被招募到启动子上，启动基因的转录。整个过程与真核生物相似，只是需要的转录因子数目要少。关于古菌的转录终止机制还知道的不多，有些古菌可能使用类似于细菌的不依赖于 ρ 因子的终止机制，因为尚未发现任何与细菌的 ρ 因子相似的终止蛋白。

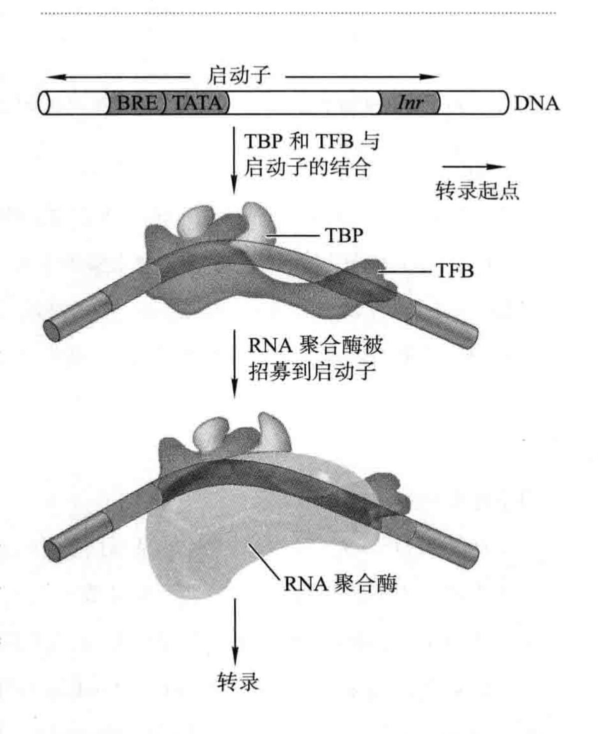


图 36-22 古菌基因转录的起始（Madigan 等，2010）

框 36-2 生化研究动态——当转录遇到 DNA 损伤

DNA 损伤随时都可能发生，损伤位点可能正好位于一个正在转录的基因内部。体外实验表明，试管中的 DNA 在转录的时候，RNA 聚合酶如果遇到缺口或空白序列（无碱基位点），可继续转录，但会漏掉损

伤部位的序列。那么，细胞内的情况又是如何呢？

如果损伤并不严重，细胞内的修复系统会及时将损伤进行修复。若损伤正好发生在一个基因的内部，并且被正在转录的RNA聚合酶遇上，还可激活专门的转录偶联修复系统，从而防止聚合酶转录出错误的RNA。如果损伤比较严重，且超过了机体修复的能力，这时不仅对DNA复制，也会对DNA转录带来不利的影响。DNA复制可以通过跨损伤合成的方式跨过损伤，但代价是引入了突变。那么，DNA转录有没有类似的跨损伤转录的方式，从而导致发生转录水平的突变呢？

为了弄清这个问题，2010年美国Emory大学的Cheryl Clauson等人设计了一个很巧妙的实验。实验以受到损伤的荧光素酶的基因为对象。之所以选择荧光素酶的基因，是因为荧光素酶可以发出容易检测的荧光。她在体外先将一个终止密码子引入到荧光素酶基因的内部，而引入的损伤就在模板链的这个位置（图36-23）。然后将带有损伤的荧光素酶基因导入到已经丧失了某些修复功能的细菌内，这实际上模拟了细菌DNA在受到严重的损伤（如辐射和大量诱变剂的作用）时的状况。此外，Clauson还使用新生霉素（novobiocin）关闭了细菌的DNA复制，这又模拟了细菌生长缓慢时的状况。

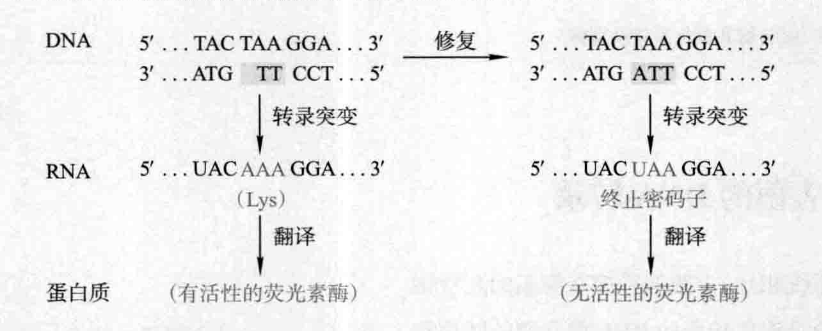


图36-23 荧光素酶基因损伤的设计及其在细胞内转录的预测

显然，只有进行跨损伤转录，且转录发生了错误，即转录没有漏掉任何序列，并将终止密码子转录成一个正常的氨基酸的密码子，才有可能产生正常的荧光素酶mRNA，并进一步翻译出有功能的荧光素。

Clauson的实验发现，损伤的荧光素酶基因在细菌体内可以通过跨损伤转录的方式，得到正常的荧光素酶，而且RNA聚合酶在缺口处可随机插入碱基，但更倾向于插入A。该结果表明，如果细胞启动了一个基因的转录，但DNA损伤还没有修复，它仍然可以导致基因在转录水平上发生突变。

科学故事——mRNA的发现

在明确DNA含有基本的遗传信息后很长的一段时间内，人们对于遗传信息是如何表达的可以说是一无所知。核苷酸序列的差异是怎样转化成一头大象和一只跳蚤之间的悬殊差别呢？即使在Watson和Crick于1953年提出DNA双螺旋结构以后的几年里，科学家对上述问题仍然是一筹莫展。

1960年，Francois Jacob和Matthew Meselson确定了蛋白质是在细胞质的核糖体上组装的。这个重要的发现表明，细胞核里的染色体和细胞质中的核糖体之间必然有一种联系的桥梁，即细胞内存在一种将细胞核里的遗传信息转移到细胞质的机制。于是，他们提出了mRNA假说。1964年，Sydney Brenner使用实验了证明了mRNA假说是正确的。

在mRNA假说提出之前，就有人曾经提出过最简单的染色体模板假说。这种假说认为，蛋白质是直接在染色体DNA上合成的。然而，该假说一提出，很快就遭到了否定，因为当时已有证据显示，蛋白质是在细胞质里合成的。那么，蛋白质究竟在细胞质的哪一个地方组装而成的呢？

为了解决这个问题，Jacob和Meselson开始选用细菌为研究对象。他们将放射性^{35}S加入到细菌的培

养基中，以使细胞内的 Cys 和 Met 能被 ^{35}S 标记，而被标记的 Cys 和 Met 会参入到合成的多肽链上，但不会参入到 DNA 或碳水化合物上。这样他们建立了一种追踪蛋白质合成的方法（图 36–24）。细菌在经过一段时间标记以后，他们使用温和的方法将它们打破，然后使用蔗糖密度梯度离心对细胞的组分进行分离，结果发现同位素标记的新合成的蛋白质与核糖体结合。如果先让 ^{35}S 短暂标记（脉冲），然后使用大量的 ^{32}S 进行追踪，就会发现与核糖体结合的放射性存留的时间很短，在追踪实验不久，就发现核糖体失去绝大多数放射性，而失去的 ^{35}S 标记出现在细胞可溶性蛋白上。这说明在脉冲标记期间合成的蛋白质已经完成并离开了核糖体。上述实验表明，蛋白质是在含有 RNA 的核糖体上从氨基酸合成而来的，一旦合成完成，就从核糖体上释放出来。

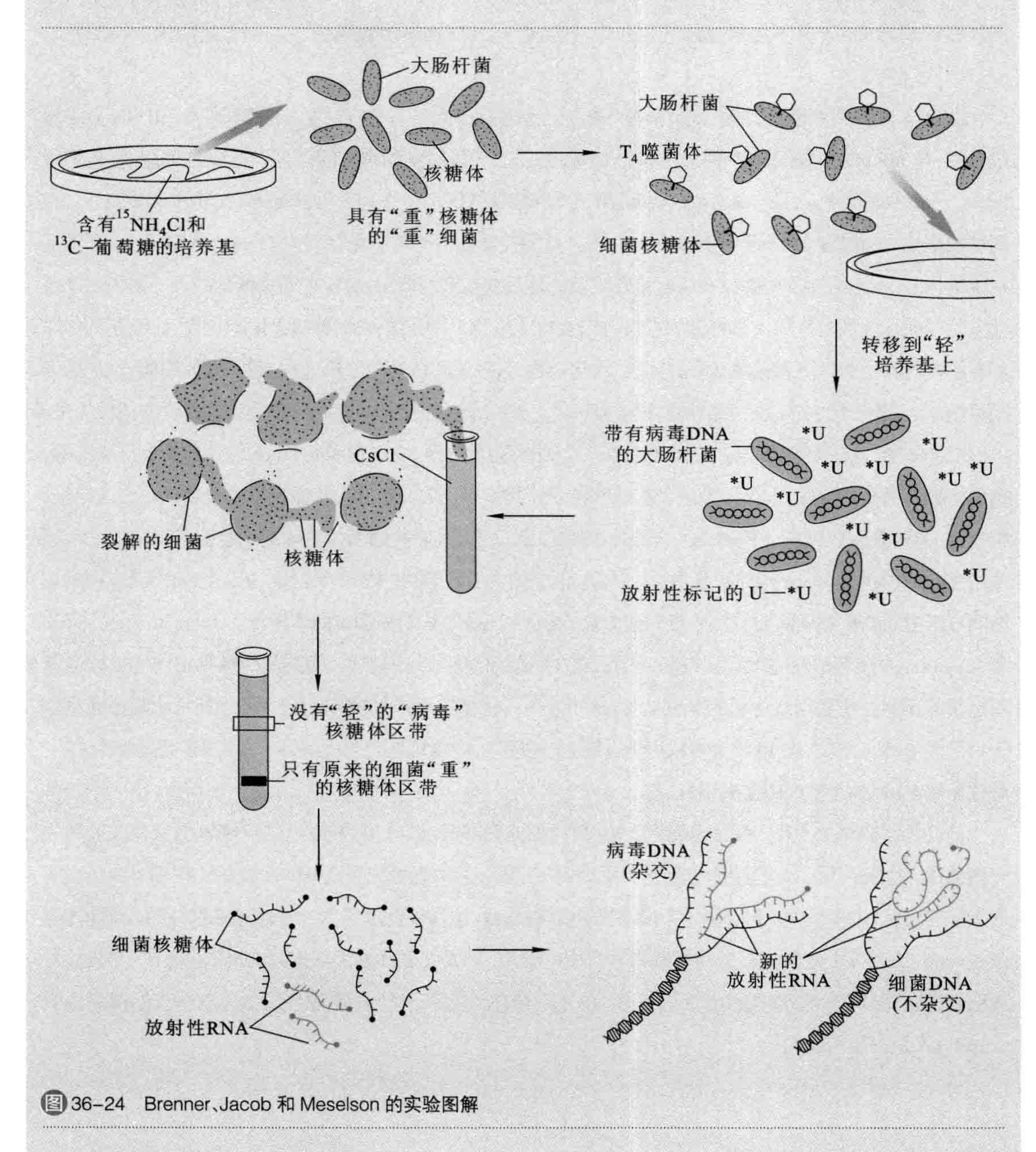


图 36–24 Brenner、Jacob 和 Meselson 的实验图解

当确定核糖体是蛋白质合成的场所以后，Jacob 和 Meselson 曾想过：会不会是核糖体里面的 RNA 作为翻译的模板呢？也许是核糖体 RNA 负责将细胞核里的遗传信息传到细胞质。如果的确如此，那细胞内就应该有许多不同的核糖体。不同的核糖体应该含有不同的 RNA 模板，而不同的 RNA 模板应该来自不同的基因，编码大小不一的蛋白质。然而，这种可能性很快就被排除，因为他们发现核糖体是完全一样的。

既然编码遗传信息的DNA在细胞核里，而遗传信息最后的表达发生在细胞质的核糖体上，而现在又发现核糖体RNA也不可能是传递信息的媒介，那么，在细胞内肯定存在其他的成分，充当遗传信息传递的载体。鉴于细胞内大多数RNA是rRNA，但并不都是rRNA，也许细胞内还有其他种类的RNA充当遗传的信使。于是在1961年，Jacob和Jacques Monod提出了信使RNA假说，认为细胞内肯定存在一种特殊的RNA是直接从DNA上合成的，它们的序列与DNA上的基因序列互补，然后被运输到细胞质为蛋白质合成提供模板。在一种蛋白质合成结束以后，其mRNA将离开核糖体，为其他的mRNA"让路"。

mRNA的假说很快得到了Brenner、Jacob和Meselson的实验确认。他们使用噬菌体感染大肠杆菌，发现感染后不久，就有一种病毒特异性的RNA被合成，并且很快和细菌内本来存在的核糖体结合。但这种新的病毒RNA并不是核糖体的永久性成分，而只与核糖体短暂结合。这不正是假说中预测到的mRNA分子吗？

mRNA假说的精髓是，细胞内应该有一类充当信使的RNA分子，它们由许多不同的mRNA分子组成，每一种mRNA在核苷酸序列上对应于DNA分子上一段特定的核苷酸序列。按照这种假说，核糖体RNA无基因特异性，这是与mRNA最重要的差别；核糖体仅仅是作为一种被动的制造蛋白质的工厂，相同的核糖体可以翻译各种不同的mRNA分子。基因特异性的遗传信息在假定的mRNA分子上，而不是在核糖体。Brenner、Jacob和Meselson接着要做的是改变基因的蓝图，检验老的核糖体能否一视同仁地合成出蛋白质。他们首先将大肠杆菌放在含有$^{15}NH_4Cl$和^{13}C-葡萄糖的培养基上培养几代，让细菌利用培养基中重的同位素氮源和碳源，作为合成碳水化合物、蛋白质和核酸的原料，使细菌上的各种成分（包括细菌核糖体）在若干代培养以后能被重的同位素标记。然后，他们使用T4噬菌体感染细菌以改变遗传信息。当时已经知道，这一类病毒感染宿主细胞以后，会破坏细菌DNA，并以自身的DNA作为遗传信息的来源，指导病毒蛋白质的合成。如果核糖体带有基因特异性的信息，那么，新的由病毒指导的核糖体也应该产生；相反，如果核糖体仅仅是一种被动的合成蛋白质的工厂，那么，老的细菌核糖体在病毒控制细菌以后照样被用来合成由病毒DNA指导的蛋白质。于是，他们的实验要解决的核心问题是，感染细菌的病毒用来制造病毒蛋白质的核糖体需要按照病毒的指令合成新的，还是原有的细菌核糖体就行。于是，Brenner、Jacob和Meselson先使用重同位素标记细菌，然后，在T4噬菌体感染细菌以后，将细菌转移到正常的轻的培养基上培养，同时，将放射线标记的RNA合成前体（^{14}C-尿嘧啶）加到培养基上。可以预计，新制造的核糖体将是轻的，新合成的RNA将会带放射性。按照上述程序，病毒指导的合成进行一段时间以后，裂解细菌，对核糖体进行CsCl密度梯度离心分析。

他们的实验结果表明，病毒感染的细胞没有制造新的轻核糖体，它们使用的核糖体仍然是以前重的旧核糖体。由此可见，T4噬菌体实际上"用旧瓶装新酒"，利用老的细菌核糖体合成新的病毒蛋白质。病毒指导的含有^{14}C标记的新RNA在病毒感染以后被合成，这种新合成的RNA先与细菌内旧的核糖体结合在一起，然后再发生解离！当将病毒指导合成的含有^{14}C标记的RNA与病毒DNA混合在一起时，发现它们很容易杂交，而将病毒RNA与细菌DNA混合在一起时，并无杂交事件发生。这些实验结果证明了mRNA假说是正确的。

e36-2 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Transcription_(genetics)（维基百科有关DNA转录的内容）
2. http://www.nature.com/scitable/topicpage/dna-transcription-426（*Nature*杂志图书馆提供的有关DNA转录的内容）
3. http://web.virginia.edu/Heidi/chapter31/chp31frameset.htm（美国弗吉尼亚大学大学生化在线课程，有关DNA转录的内容）

4. http://bioweb.wku.edu/courses/biol22000/18TranscInitProk/default.html（美国西肯塔基大学基础分子生物学课程，有关原核基因转录的内容）
5. http://bioweb.wku.edu/courses/biol22000/19TranscInitEuk/default.html（美国西肯塔基大学基础分子生物学课程，有关真核基因转录的内容）

参考文献：

1. Ringel R, Temiakov D, et al. Structure of human mitochondrial RNA polymerase. *Nature*, 478 (7 368): 269–273, 2011.
2. Soultanas P, et al. Co-directional replication-transcription conflicts lead to replication restart. *Nature*, 470: 554–557, 2011.
3. Clauson C L, et al. Abasic sites and strand breaks in DNA cause transcriptional mutagenesis in *Escherichia coli*. *Proc Natl Acad Sci*, 107 (8): 3 657–3 662, 2010.
4. Svetlov V, Nudler E. Clamping the clamp of RNA polymerase. *The EMBO Journal*, 30: 1 190–1 191, 2011.
5. Woychik N A, Hampsey M. The RNA polymerase Ⅱ machinery: structure illuminates function. *Cell*, 108: 453–463, 2002.
6. Gross C A, et al. Views of transcription initiation. *Cell*, 109: 417–420, 2002.
7. Conaway R C, et al. Control of elongation by RNA polymerase Ⅱ. *Trends Biochem Sci*, 25: 375–380, 2000.
8. Kornberg R D. RNA polymerase Ⅱ transcription control. *Trends Biochem Sci*, 21: 325–327, 1996.

第三十七章　转录后加工

基因转录的直接产物称为初级转录物（primary transcript）。初级转录物一般是无功能的，它们在细胞内必须经历一些结构和化学的变化，即转录后加工（post-transcriptional processing）以后，才会有功能。转录后加工可能是 RNA 的功能所必需的，也可能是基因表达调控的一种手段。

RNA 所能经历的后加工方式可达 10 种以上，但后加工反应的本质要么是增减一些核苷酸序列，要么是修饰某些特定的核苷酸。

在细菌、古菌和真核生物中，三种主要的 RNA（mRNA、rRNA 和 tRNA）所经历的后加工反应并非完全相同，即使同一种 RNA 前体也可能有不同的加工路线，后一种情形可导致一个基因产生几种终产物，这已成为真核生物基因表达调控的一种很重要的手段。

本章将分别介绍 mRNA、rRNA 和 tRNA 在细菌、真核生物和古菌中所经历的各种后加工反应。

第一节　细菌的转录后加工

一、mRNA 前体的后加工

细菌的 mRNA 一般很不稳定，很少经历后加工。它们通常在几分钟内完成转录、翻译和降解。实际上，绝大多数 mRNA 在 3′端还没有被转录之前，核糖体就结合到 5′端开始翻译，而当翻译到 3′端的时候，5′端又开始降解。虽然有些细菌的 mRNA 可在 3′端引入多聚 A 尾巴，但这种尾巴结构也是充当 mRNA 降解的信号。此外，某些噬菌体 mRNA 会发生最简单的剪切反应，将一个多顺反子切割成单顺反子，也有某些噬菌体的 mRNA 含有内含子，需要经过剪接才能成熟。

二、rRNA 前体的后加工

细菌和真核生物的 rRNA 基因都是以多顺反子的形式存在，而在细菌中，还可能带有某些 tRNA 基因。要从这种共转录物中分别得到 3 种 rRNA，首先需要剪切，使它们从共转录物中释放出来，然后还要进行修剪，以去除两端多余的核苷酸序列，此外，三种 rRNA 还需要进行一些特定的化学修饰。因此，细菌 rRNA 的后加工包括剪切（cleavage）、修剪（trimming）和核苷酸的修饰（图 37–1）。

(1) 剪切和修剪　由特定的核糖核酸酶催化，是在 rRNA 与核糖体蛋白结合以后发生的，涉及的酶包括核糖核酸酶Ⅲ、D、F、P、M16 和 M5 等。其中剪切为“粗加工”，由内切酶从共转录物的内部将各 rRNA 两侧的多数核苷酸切除，并将单独的 rRNA 释放出来；修剪为“细加工”，由外切酶从各 rRNA 两端逐个水解掉多余的核苷酸。在加工过程中，16S 和 23S rRNA 前体之间通过碱基互补配对形成茎环结构，然后才能成为核糖核酸酶Ⅲ的底物。核糖核酸酶Ⅲ在茎部切开 rRNA 前体，释放出单体。

(2) 核苷酸的修饰　发生在剪切和修剪反应之前，其主要形式为核糖 2′–OH 的甲基化，甲基供体为 SAM。甲基化的功能是消除 2′–OH 的亲核性，从而提高 rRNA 的稳定性。

三、tRNA 前体的后加工

细菌的 tRNA 前体的后加工方式也包括剪切和修剪以及核苷酸的修饰（图 37–2）。

(1) 剪切和修剪　一个典型的成熟 tRNA 约由 80 nt 组成，但其前体可能有几百个核苷酸。多出

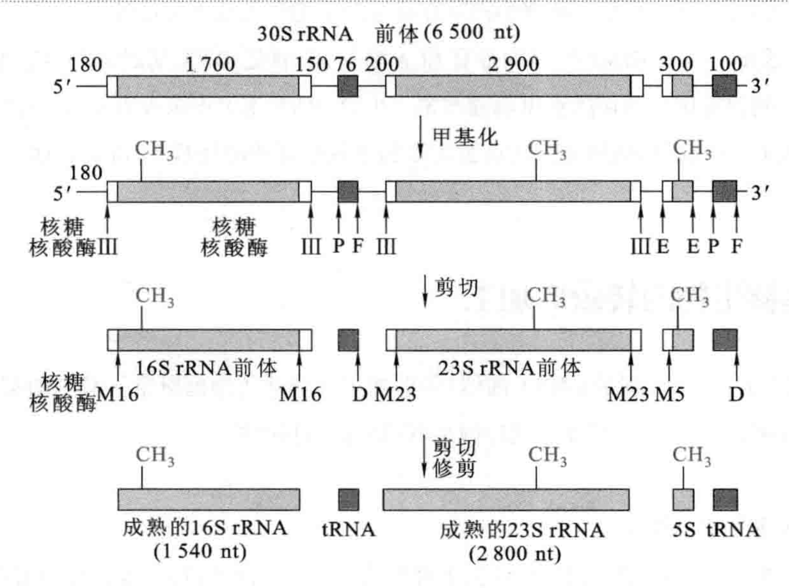


图37-1 细菌 rRNA 前体的后加工

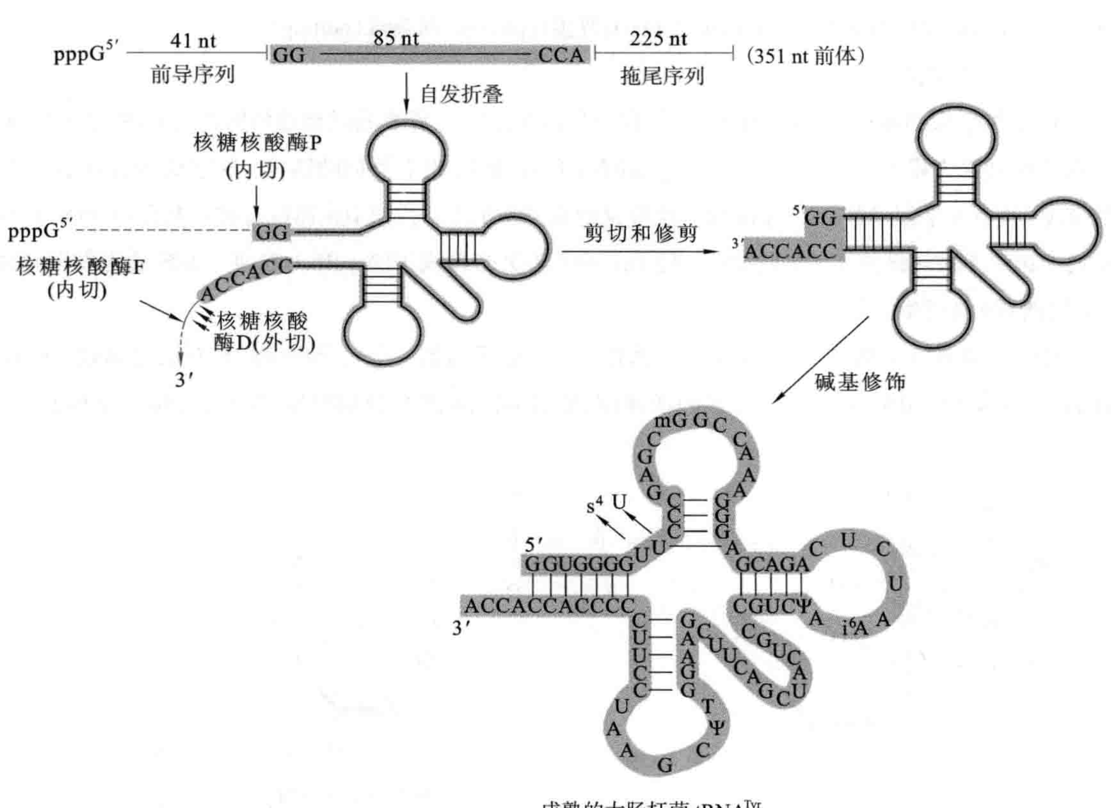


图37-2 大肠杆菌 tRNA^{Tyr} 的后加工

来的核苷酸一般分布在前体的两侧,其中位于5′端的称为前导序列(leader sequence),而位于3′端的称为拖尾序列(trailer sequence)。无论是前导序列,还是拖尾序列,都需要被特定的核酸酶切除。某些tRNA 基因单独转录,某些与 rRNA 基因组成多顺反子共同转录。在混合型 tRNA-rRNA 前体中,内部的 tRNA 在受到特定的核糖核酸酶作用以后,得以释放。

在各种核酸酶中,有两种最重要:一种是参与5′端剪切的核糖核酸酶 P,它已被证明是一种名副其实的核酶(参看第十一章“核酶的结构与功能”),另一种是参与3′端剪切的核糖核酸酶 Z。核糖核酸酶 Z 负责切除基因内部缺乏 CCA 序列的 tRNA 的拖尾序列,切点在拖尾序列中第一个没有配对的核苷酸的3′端。如果一种 tRNA 的基因本来就有 CCA 序列,就由核糖核酸酶 F 和 D 切除拖尾序列。

(2) 核苷酸的修饰　一个成熟的 tRNA 分子约有 10% 的核苷酸是被修饰的。修饰主要集中在碱

基上，修饰的方式有近百种，而每一种修饰都是在特定的修饰酶催化下完成的。

(3) CCA的添加　所有tRNA的3′端都有CCA序列，但细菌tRNA基因的大多数本来就有CCA，故只需要将其右侧的拖尾序列切除就可暴露出来。少数tRNA基因本来没有CCA，就需要在tRNA核苷酸转移酶催化下，在核糖核酸酶Z剪切后留下来的3′端核苷酸的羟基上，依次转移上CCA序列。

第二节　真核生物的转录后加工

真核生物RNA前体(特别是mRNA前体)的后加工反应远比细菌复杂。其中的某些后加工反应的性质与细菌很相似，而另一些后加工反应则是真核生物所特有的。

一、mRNA前体的后加工

Quiz1 真核生物细胞核mRNA最早发生的后加工形式是哪一种?

真核生物的核mRNA必须在细胞核经历多种形式的后加工反应以后，才会成为有功能的分子，也并被运输出细胞核，再在细胞质作为翻译的模板。这些后加工反应包括：5′端"戴帽"(capping)、3′端"加尾"(tailing)、内部甲基化(internal methylation)、剪接(splicing)和编辑(editing)。

(一) 5′端"戴帽"

在真核生物细胞核mRNA的5′端，含有一个以5′,5′-三磷酸酯键相连的修饰鸟苷酸，这种特殊的鸟苷酸被称为帽子结构(图37-3)，某些snRNA和snoRNA也有类似的结构。帽子结构有0型、1型和2型：具有0型帽子的mRNA前两个被转录的核苷酸的2′-核糖OH都没有被甲基化；1型帽子的mRNA第一个被转录的核苷酸的2′-核糖OH被甲基化；2型帽子的mRNA前两个被转录的核苷酸的2′-核糖OH都发生了甲基化。

帽子一般有4个功能：①提高mRNA的稳定性。帽子与第一个转录出来的核苷酸之间独特的连接方式，可保护mRNA抵抗5′-外切酶的水解，原核mRNA的半衰期极短可能与无帽子结构有关。

图37-3　真核细胞mRNA的帽子结构

②参与识别起始密码子的过程，提高 mRNA 的可翻译性。这也就解释了为什么几乎所有的真核生物的细胞核 mRNA 都是单顺反子。③有助于 mRNA 被运输到细胞质。④提高剪接反应的效率。

加帽反应是一种共转录反应，一般在转录物的5′端从RNA聚合酶离开通道内暴露出来以后就开始了。但为什么 tRNA 和 rRNA 没有帽子结构呢？之所以只有 mRNA 以及某些 snRNA 和 snoRNA 才有帽子结构，是因为它们都由 RNA 聚合酶Ⅱ催化。事实上，任何由 RNA 聚合酶Ⅱ催化转录的 RNA 肯定都有帽子结构，除非后来被剪切掉了。在转录进入延伸阶段以后，TFⅡH 很快磷酸化 CTD 重复序列中的 Ser5。磷酸化的 Ser5 会将转录因子 DSIF 招募到转录复合物。而 DSIF 随后又将另一种转录因子 NELF 招募进来，使转录暂停。上述暂停允许加帽酶进入，修饰转录物的5′端。第三种转录因子 P-TEFb 是一种激酶，在帽子结构形成不久，也被招募到复合物，然后磷酸化 CTD 的 Ser2 和 NELF。NELF 随之失活，转录得以继续延伸（图 37-4）。

Quiz2 若是让 RNA 聚合酶Ⅱ催化 rRNA 和 tRNA 基因的转录，那么转录出来的 rRNA 和 tRNA 有无帽子结构？为什么？

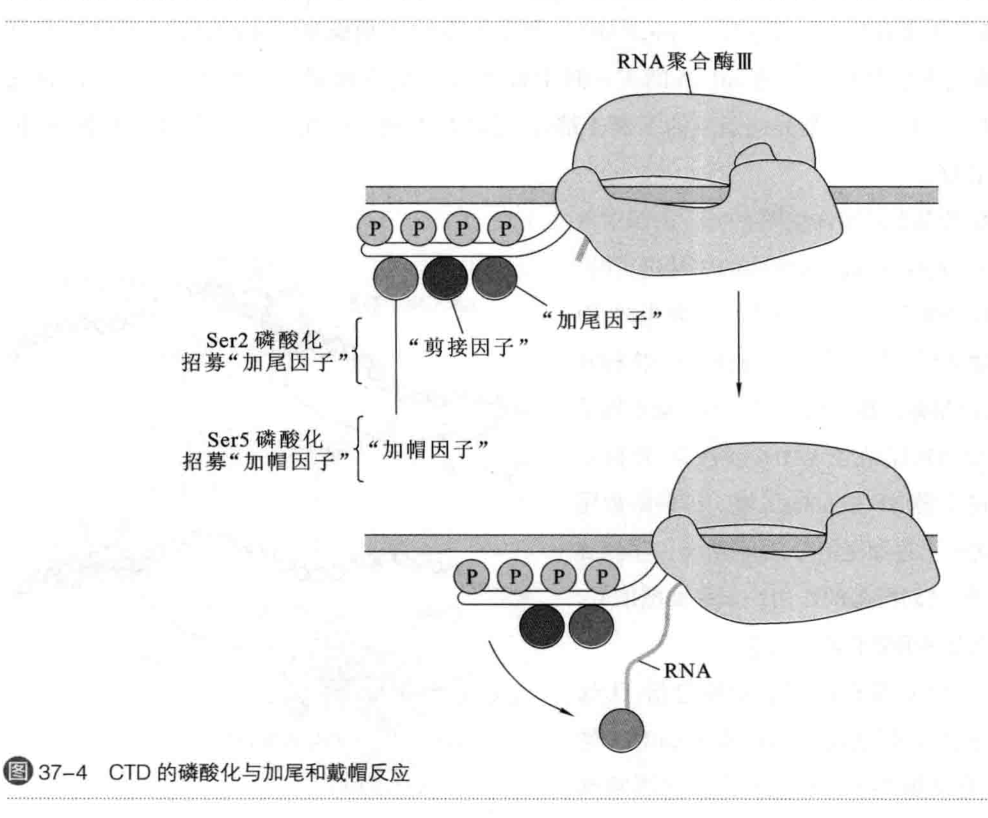


图 37-4　CTD 的磷酸化与加尾和戴帽反应

（二）3′端"加尾"

多聚 A 尾巴可视为大多数真核细胞核 mRNA 3′端特有的结构，其长度在 250 个 A 左右。在少数细菌、古菌和哺乳动物的线粒体以及植物的叶绿体或线粒体中，虽然也有 mRNA 含有多聚 A 尾巴，但与真核生物细胞核 mRNA 的尾巴相比，尾巴较短、不稳定，且功能不一样。在细菌和植物的叶绿体或线粒体中，其中的 mRNA 的尾巴是多聚核苷酸磷酸化酶（PNP）进行降解的信号，而哺乳动物的线粒体 mRNA 可通过加尾，在3′端引入终止密码子或者调节稳定性。由于含有多聚 A 尾巴的 mRNA 的编码链上并无多聚 A 序列，因此多聚 A 尾巴是在转录后添加上去的。

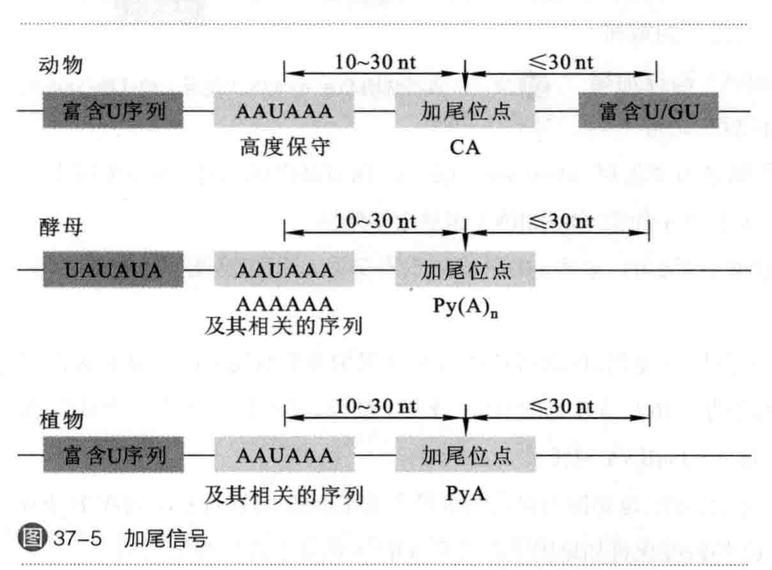


图 37-5　加尾信号

加尾反应是精确的，由两种因素控制：一种为顺式元件（图 37-5），位于 mRNA 前体内

部，为特殊的核苷酸序列，充当加尾信号，可将其视为加尾反应的“内因”；另一种为识别加尾信号的蛋白质或催化加尾反应的酶，可将其视为加尾反应的“外因”。

最重要的加尾信号是 AAUAAA。此外，还可能有其他次要的加尾信号。例如，动物在加尾点的下游有一段富含 U/GU 的序列，还可能在 AAUAAA 的上游有一段富含 U 的序列。再如，酵母在 AAUAAA 序列的上游还有一段 UAUAUA 序列，而植物 AAUAAA 信号的上游有一段富含 U 的序列。

参与加尾反应的“外因”有：①剪切 / 多聚腺苷酸化特异性因子（cleavage and polyadenylation specificity factor，CPSF），此蛋白质识别和结合 AAUAAA 序列，并参与和多聚 A 聚合酶以及剪切刺激因子的相互作用；②剪切刺激因子（cleavage stimulation factor，CstF），此蛋白质的功能是识别 GU/U 序列并与 CPSF 结合，刺激剪切反应；③剪切因子Ⅰ和Ⅱ（cleavage factor Ⅰ/Ⅱ，CF Ⅰ/Ⅱ），为特殊的内切核酸酶，负责剪切反应；④多聚 A 聚合酶（polyA polymerase，PAP），一种不需要 DNA 模板并只对 ATP 有亲和性的 RNA 聚合酶，此酶负责催化在切开的 mRNA 的 3′-OH 上添加多聚 A；⑤多聚 A 结合蛋白（polyA binding protein，PABP），其主要的功能是与新生的多聚 A 结合，提高 PAP 的进行性，刺激多聚 A 的延伸，并控制多聚 A 的长度。

加尾反应的基本步骤详见图 37-6。由图中可见，加尾并不是直接在 mRNA 前体的 3′端进行的，而是在末端核苷酸上游某一特定的位置发生的，因此在真正的加尾反应发生之前，先要进行一次精确的剪切反应，以暴露出加尾处的 3′-OH。至于加尾点的选择主要由加尾信号 AAUAAA 决定，该信号位于加尾点的上游 20~30 nt 的区域。正确的加尾反应不仅取决于上述加尾信号，还取决于一系列酶和蛋白质因子，它们所起的作用是识别加尾信号，并催化剪切以及多聚腺苷酸化反应。

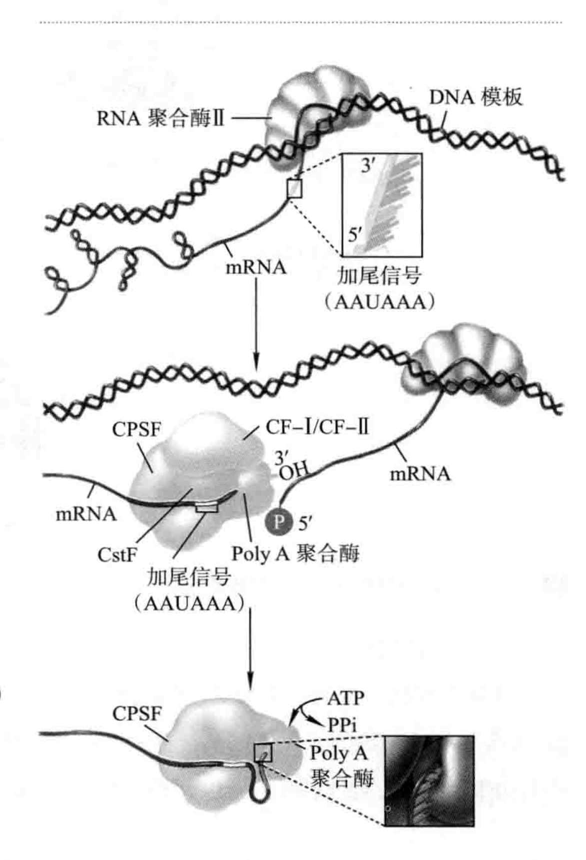


图 37-6　真核细胞 Pre-mRNA 加尾反应模式图（Robert Horton，2006）

多聚 A 尾巴所具有的可能功能包括：①保护 mRNA 抵抗 3′- 外切酶的消化，提高 mRNA 的稳定性。已有证据表明，尾巴的长度能影响到 mRNA 的稳定性，而且这可能是受到调控的。当尾巴的长度降到一定的临界值(酵母少于 10~15 A)以下，常常充当 mRNA 的降解信号。② PABP 能够与帽子相互作用，增强 mRNA 的可翻译性，提高其翻译的效率。③影响最后一个内含子的剪接。④某些本来缺乏终止密码子的 mRNA 通过加尾反应创造终止密码子。例如在 UG 后加尾可产生 UGA，在 UA 后加尾产生 UAA。⑤通过可变加尾（alternative tailing）调节基因的表达（参看第四十三章“真核生物的基因表达调控”）。⑥有助于加工好的 mRNA 运输出细胞核。

多聚A尾巴并不是mRNA翻译必不可少的。某些mRNA虽然没有多聚A，但仍然能够正常地翻译，比如组蛋白的mRNA。

Quiz3　组蛋白的 mRNA 既不能进行剪接，又不能进行加尾，你认为其中的原因和意义何在？

由于真核细胞只有 mRNA 才有多聚 A 尾巴，因此可使用含有寡聚胸苷酸（oligo T）或寡聚尿苷酸（oligo U）的树脂，将带有多聚 A 尾巴的 mRNA 与其他的 RNA 分开。此外，还可以使用人工合成的寡聚脱氧胸苷酸作为引物，将含有多聚 A 的 mRNA 逆转录成 cDNA。

与加帽反应一样，只有 mRNA 才会加尾，也是因为聚合酶Ⅱ最大亚基上的 CTD 重复序列被 TFⅡH 磷酸化，但磷酸化位点为 Ser2。Ser2 的磷酸化将加尾因子招募到 mRNA 前体上进行加尾反应。

（三）内部甲基化

mRNA 的内部甲基化主要是指 mRNA 分子上的某些腺嘌呤（通常在 GAC 序列之中）N6 经历甲基化修饰，形成 6- 甲基腺嘌呤（N6-methyladenosine，m^6A）的过程。内含子和外显子上都可能发生这样的修饰。这种修饰的功能尚不特别清楚，但有证据表明，催化 A 发生甲基化修饰的甲基转移酶是拟南芥胚胎发育所必需的。

（四）剪接

剪接这种后加工方式是在发现基因断裂现象后确定的。1977 年，由 Phillip Sharp 和 Richard Roberts 领导的两个实验小组几乎同时在腺病毒的晚期表达基因中，发现蛋白质基因断裂现象。

他们使用的是 R 环技术，其操作流程如图 37-7 所示，将一种蛋白质基因的 mRNA 与变性的基因组 DNA 进行杂交，如果其基因是连续的，那 mRNA 的编码区序列与其模板链序列呈连续互补，就不可能有非互补序列以环的形式突出来；反之，如果基因是断裂的，那在杂交以后，由于基因编码链上的部分序列已不存在于成熟的 mRNA 分子上，就会以 R 环的形式凸现出来，在电镜下可以观测到。由于 mRNA 容易水解，也可以先将它逆转录成为稳定的 cDNA 后，代替 mRNA 进行杂交实验。

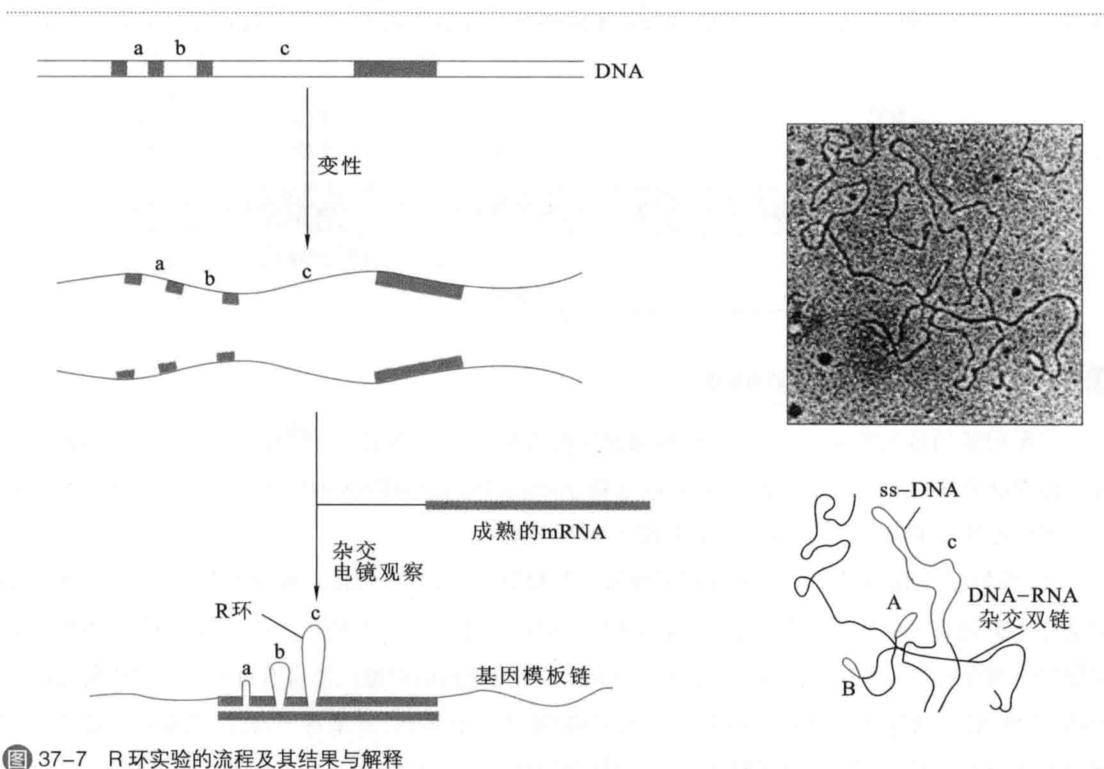


图 37-7　R 环实验的流程及其结果与解释

Sharp 和 Roberts 当时的实验结果出现了 3 个 R 环，这表明他们研究的基因断裂了 3 次。到了 1978 年上半年，人们使用 R 环技术相继发现很多蛋白质以及某些 tRNA 和 rRNA 基因也是不连续的。同年，Walter Gilbert 引入了外显子（exon）和内含子（intron）这两个概念，分别表示蛋白质基因中编码氨基酸和不编码氨基酸的碱基序列。一个断裂基因在转录的时候，外显子和内含子一起被转录在同一个初级转录物之中，但内含子并不出现在最终成熟的 mRNA 分子上，而是被剪切出去了。细胞内这种将内含子去除并将相邻的外显子连接起来的过程称为剪接。

进一步研究表明，基因断裂普遍存在于真核细胞及其病毒的基因组中，在高等生物的基因组中，只有少数蛋白质基因是连续的（如组蛋白和干扰素），但在低等的真核生物中，断裂基因却不多见，例如，酵母基因组的大约 6 000 个基因中，含有内含子的基因仅有 239 个，而且通常只有 1 个内含子。一般说来，一个典型的真核生物蛋白质的基因由 10% 的外显子序列和 90% 的内含子序列组成。

不同断裂基因含有的内含子数目不一定相同，同样内含子大小也会有差别。例如，鸡卵清蛋白的基因长达 7.7×10^3 nt，有 6 个内含子，经过剪接，最后成熟的 mRNA 只有 1 872 nt。而抗肌营养不良蛋

Quiz4 人类基因组中，最大的内含子有多大？内含子数目最多的基因有多少个内含子？

白的基因（dystrophin）长度在 2×10^6 nt 以上，共有 78 个内含子，而成熟的 mRNA 只有 1.4×10^4 nt！

Sharp 和 Roberts 的发现从根本上改变了基因是连续的这一传统观念，在分子生物学发展过程中具有划时代的意义，他们因此获得了 1993 年的诺贝尔生理学或医学奖，然而内含子是如何被切除的呢？

1. 真核细胞 mRNA 前体剪接的机理

mRNA 前体的剪接是高度精确的。其精确性一方面取决于位于外显子和内含子交界处的剪接信号——“内因”，另外一方面取决于 5 种被称为 snRNP 的核糖核酸蛋白质复合物——“外因”。

（1）剪接反应的“内因” 通过比较多种 mRNA 剪接点周围的碱基序列后发现，外显子和内含子交界处存在高度保守的一致序列，这些序列充当控制剪接的顺式元件（图 37–8）。从图中可以看出，作为顺式元件的保守序列存在于直接参与剪接反应的 5′– 剪接点（又名供体位点或左剪接点）、3′– 剪接点（又名受体位点或右剪接点）和分支点(branch point)。5′– 剪接点的保守序列是内含子的前两个 GU，3′– 剪接点的保守序列是内含子的最后两个 AG，此规律有时称为“GU–AG 规则”。位于内含子内部的分支点序列的保守性不高，但酵母细胞的分支点一致序列几乎总是 UACUAAC，该序列中的一个 A 对于剪接反应的发生是至关重要的。此外，在 3′– 剪接点的上游不远处，还有一段主要由 11 个嘧啶组成的序列。另外，很多外显子内还有能增强剪接效率的外显子剪接增强子（exonic splicing enhancer，ESE）。

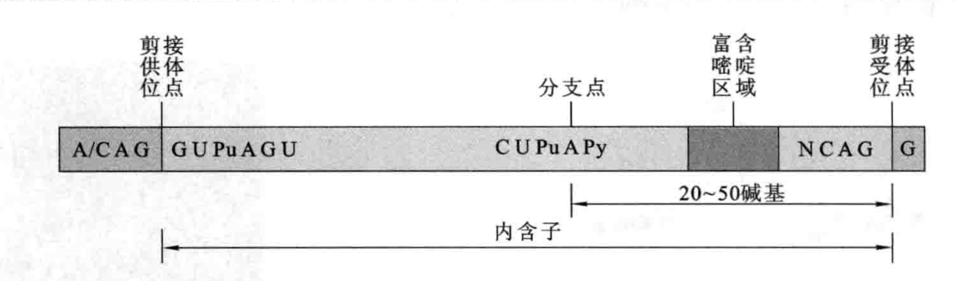


图 37–8 真核细胞核 Pre–mRNA 的剪接信号

如果剪接信号发生突变，不仅会影响到剪接的精确性，还可能影响到剪接的效率。已发现，至少有一种类型的 β– 地中海贫血（thalassemia），是由于 β– 珠蛋白的 mRNA 前体在外显子和内含子交界处的剪接信号发生了突变，导致内含子不能被有效地去除。

（2）剪接反应的“外因” 为了保证剪接反应的精确进行，光有剪接信号还不行，细胞还必须具备识别这些剪接信号的机制。此外，在剪接反应中，相邻的外显子需要被并置在一起。那么，细胞是如何做到这两点呢？ snRNP 的发现为此提供了线索。真核细胞的细胞核里含有许多序列高度保守的小 RNA，这些 RNA 统称为 snRNA。snRNA 一般是由 60~300 nt 组成，且富含 U，它们通常与一些特殊的蛋白质结合在一起，形成核小核糖核酸蛋白质复合物（small nuclear ribonucleoprotein，snRNP）。

已发现 10 多种类型的 snRNP，但参与剪接的只有 U1、U2、U4、U5 和 U6（表 37–1）。其中 U1 含有与 5′– 剪接点互补的序列，因此它的功能是识别 5′– 剪接点的信号；U2 含有与分支点互补的序列，因此它的主要功能是识别分支点。此外，U2 与 U6 之间也存在互补碱基对，它们之间通过配对形成的两段双螺旋对于剪接反应非常重要；U5 虽然缺乏任何与剪接信号互补的序列，但它能够与一个外显子的最后一个核苷酸以及下一个外显子的第一个核苷酸结合，这种结合似乎有助于将两个相邻的外显子

表 37–1 参与剪接反应的 5 种 snRNA

snRNA	互补性	功能
U1	内含子的 5′端	识别和结合 5′– 剪接点
U2	分支点	识别和结合分支点。在剪接体组装中，也与 U6 snRNA 配对
U4	U6 snRNA	结合并失活 U6。在剪接体组装中，U4–U6 之间的碱基配对被 U2–U6 之间的剪接配对所取代。U6 也取代 U1 与 5′– 剪接点的相互作用
U5	相邻的外显子	与相邻的 2 个外显子结合，以防止它们离开剪接体
U6	U4 和 U2	在剪接体中与 U2 结合，参与催化

拉到一起；U6 不仅能与内含子的 5′端互补配对，还能够与 U4 之间配对形成紧密的复合物。

除了 snRNP 之中的蛋白质成分参与剪接作用以外，细胞内还有一些游离的蛋白质因子也与剪辑有关，它们称为剪接因子（splicing factor）。剪接因子中最重要的一类是富含 Ser 和 Arg 的蛋白（serine/arginine-rich protein，SR 蛋白），它们所起的作用主要是与 ESE 结合，调节可变剪接。

(3) 剪接体（spliceosome） 主要是由 mRNA 前体、mRNA 前体结合蛋白质和 5 种 snRNPs 在细胞核内按照一定的次序组装起来的超分子复合物，其大小相当于核糖体。剪接反应是在剪接体内发生的。剪接反应本身并不复杂，由 2 次连续的转酯反应组成（图 37–9）：第一次转酯反应发生在 5′– 剪接点，首先由分支点内的一个腺苷酸残基上的 2′–OH 对 5′– 剪接点处的磷酸二酯键做亲核进攻，而导致该位点的 3′，5′– 磷酸二酯键发生断裂，在 5′– 剪接点上的磷酸二酯键断裂的同时，分支点腺苷酸 2′–OH 与内含子 5′端磷酸之间形成 2′，5′– 磷酸二酯键；第二次转酯反应发生在 3′– 剪接点，由刚刚从 5′– 剪接点游离出来的外显子 3′–OH 对 3′– 剪接点的磷酸二酯键做亲核进攻，致使内含子以套索（lariat）结构的形式释放，同时，两个相邻的外显子通过新的磷酸二酯键连接起来。

Quiz5 细胞内的哪些过程利用转酯反应？

(4) 剪接体的组装 剪接体的组装是一个有序、耗能的过程（图 37–10），而且它本身也处在动态变化之中，随着剪接反应的进行，某些成分可能离开或进入剪接体。一旦剪接反应完成，剪接体即发生解体。

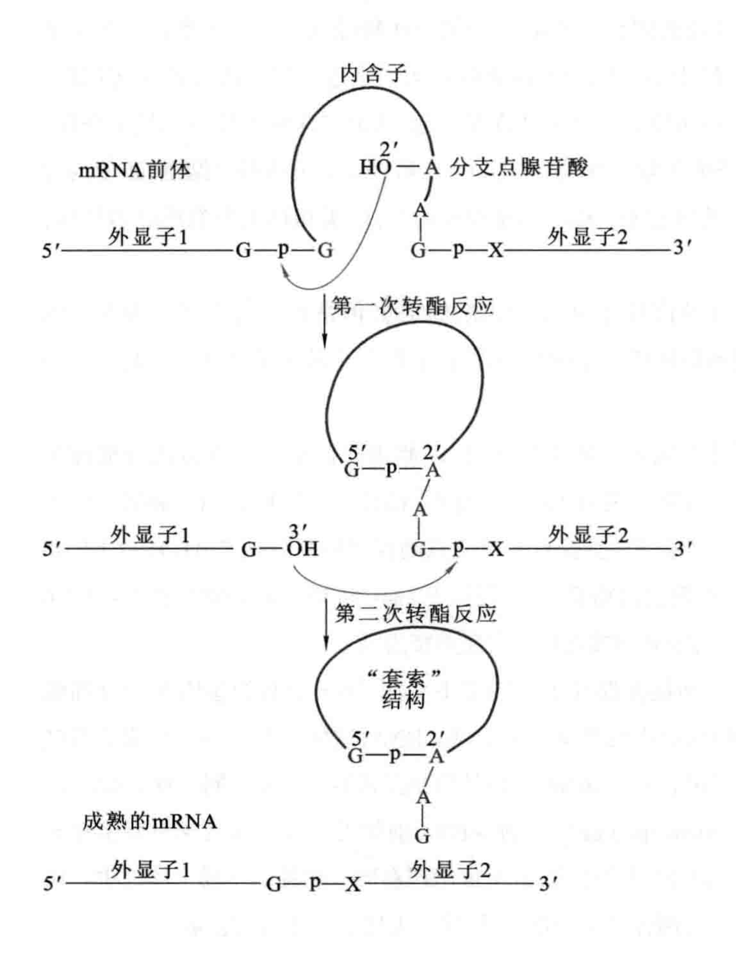


图 37–9 剪接的两次转酯反应

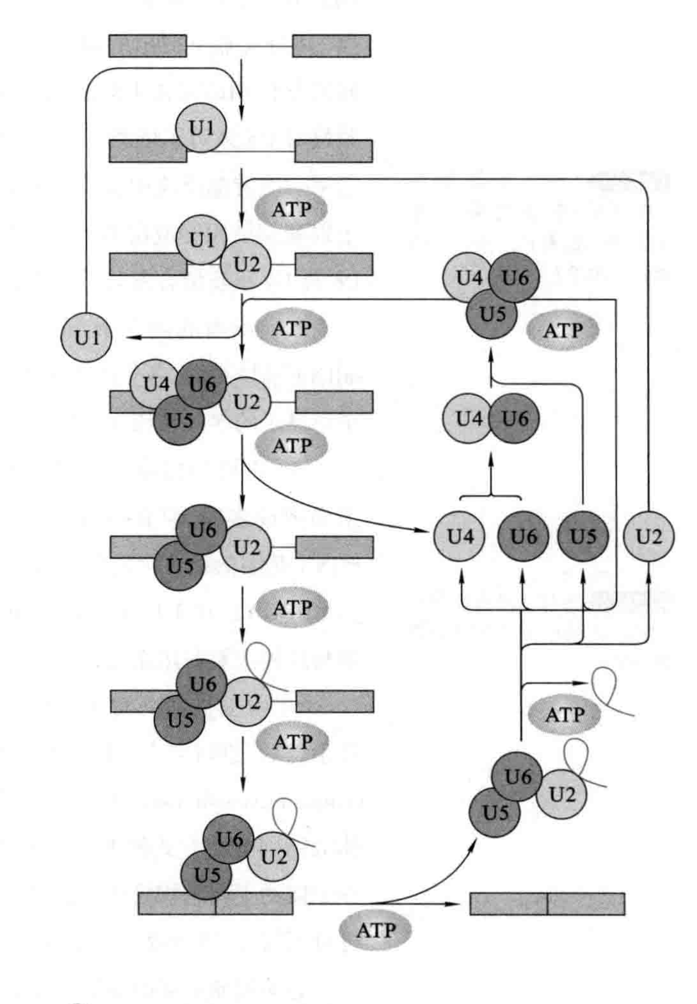


图 37–10 剪接体的组装和去组装

剪接体组装的大致过程依次是：①分支点结合蛋白（branch point binding protein，BBP）识别分支点并与它结合，剪接因子 U2AF 与富含嘧啶的序列结合。②U2 取代 BBP，并与分支点周围的一致序列配对。然而，U2–snRNA 与分支点一致序列之间的配对并不完美，在分支点内唯独有 1 个 A 无互补配对的 U，因此，这个“孤独”的 A 就突出在双螺旋外而被激活，为剪接的第一次转酯反应提供了便利（图 37–11）。

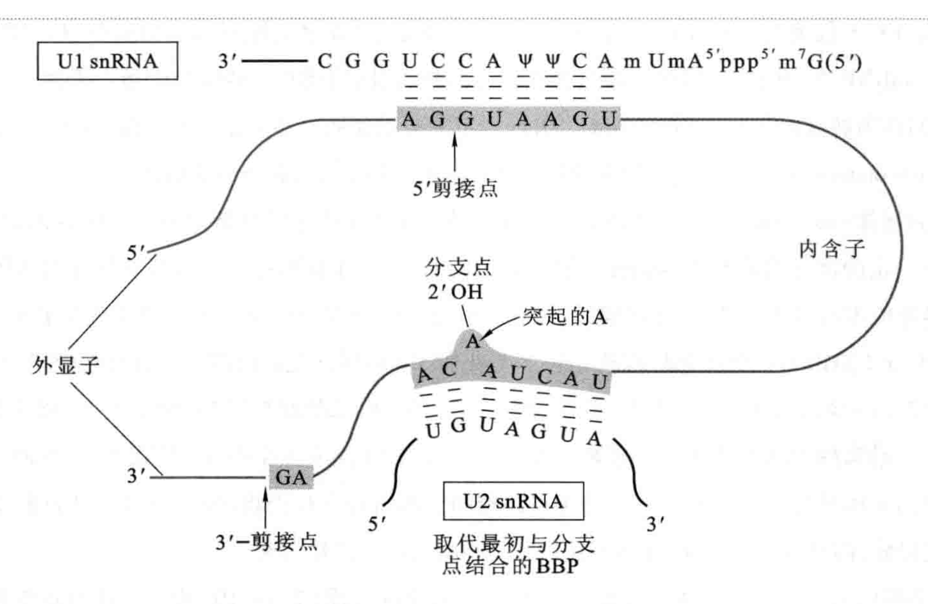


图 37-11 分支点腺苷酸 2′-羟基的突出以及 5′-剪接点和 3′-剪接点的结构

③U1-snRNP 与 5′-剪接点的一致序列互补配对。④ U4/U6-U5 三聚体进入剪接体。⑤ U6 一方面代替 U1 与 5′-剪接点的一致序列结合，另一方面与 U4 脱离，转而与 U2 结合。⑥ U6 的"脚踩两只船"将分支点上突出的腺苷酸的 2′-OH 拉近到 5′-剪接点，为第一次转酯反应创造了条件。在这里起催化作用的是由 U6 和 U2 共建的活性中心，其中 U6 的贡献更大。在进行第一次转酯反应以后，紧接着 U5 经历了依赖于 ATP 的重排，将相邻的外显子并置在一起，为第二次转酯反应创造了条件。⑦第一次转酯反应中游离出来的外显子的 3′端 -OH 亲核进攻 3′-剪接点上的磷酸二酯键，致使内含子以套索结构释放出来，而外显子则被连接起来。⑧一旦剪接反应完成，剪接体的所有成分即解体，U6 与 U4 重新结合参与下一轮剪接反应。

Quiz6 为什么细胞核 mRNA 的剪接需要消耗 ATP，而 I 组和 II 组内含子的剪接不需要？

虽然所有的内含子 / 外显子交界处的信号序列几乎都是相同的，但是对于含有多个内含子的 mRNA 前体来说，各个内含子的去除仍然是按照一定的次序进行的，前一个外显子的 3′端总是与后一个外显子的 5′端精确地连接起来。

（5）次要剪接途径　GU-AG 规则适合绝大多数断裂基因，以此为剪接信号的剪接途径被称为主要剪接途径，但有些内含子的剪接信号并不遵守 GU-AG 规则，而是以 AU 开头，AC 结尾。如人 PCNA 基因的第 6 个内含子，这一类内含子在 5′-剪接点和分支点的保守序列分别是 AUAUCCUY 和 UCCUURAY，由 U11 和 U12-snRNAs 分别跟它们结合，另外还有 U4atac 和 U6atac 分别代替 U4 和 U6 参与这种次要剪接途径，只有 U5 同时参与主要剪接途径和次要剪接途径。

Quiz7 为什么真核细胞核 mRNA 的两条剪接途径都需要 U5？

（6）可变剪接　对于一些基因来说，剪接套路几乎是固定不变的：同一个断裂基因在所有细胞转录以后，进行完全相同的剪接反应，最后产生的是单一的成熟 mRNA，这样的剪接称为组成型剪接（constitutive splicing）。但对于多数基因来说，一个 mRNA 前体的剪接方式不止一种。同一种 mRNA 前体的不同剪接方式称为可变剪接（alternative splicing）。一种 mRNA 前体的可变剪接现象可发生在同一个体的不同组织之中，或者同一个体的不同发育阶段，甚至可出现在同一物种的不同个体之中。通过可变剪接，可导致一个基因编码出两种或两种以上的蛋白质，这大大增加了基因的容量。

有关可变剪接的实例可参看"真核生物基因表达的调控"。mRNA 前体之所以能够发生可变剪接，一方面是因为它含有多重剪接信号，另一方面是因为细胞中存在一些调节剪接反应的蛋白质。

（7）反式剪接　前面介绍的剪接反应都发生在同一个 RNA 前体上，因此称为顺式剪接（*cis*-splicing）。如果剪接发生在两种不同的 RNA 前体分子之间，就称为反式剪接（*trans*-splicing）。反式剪接仅发生在由某些叶绿体基因组和某些生物（如锥体虫和线虫）的核基因组编码的 mRNA 上。例如，锥体虫通过反式剪接，将相同的由 35 nt 组成的前导序列从一种由 RNA 聚合酶Ⅱ催化转录的 RNA 分子

上，加到所有的由 RNA 聚合酶Ⅰ催化转录的 mRNA 的 5′端，从而使它们也能有帽子结构。再如，秀丽线虫（*C. elegans*）很多基团形成类似于原核生物的操纵子结构，转录产生的是无翻译活性的多顺反子，通过反式剪接可以将多顺反子裂解成单顺反子，同时让由 22 nt 组成的带有帽子结构的前导序列，添加到从多顺反子中游离出来的缺乏帽子结构的单顺反子的 5′端，使它们也具有翻译的活性。

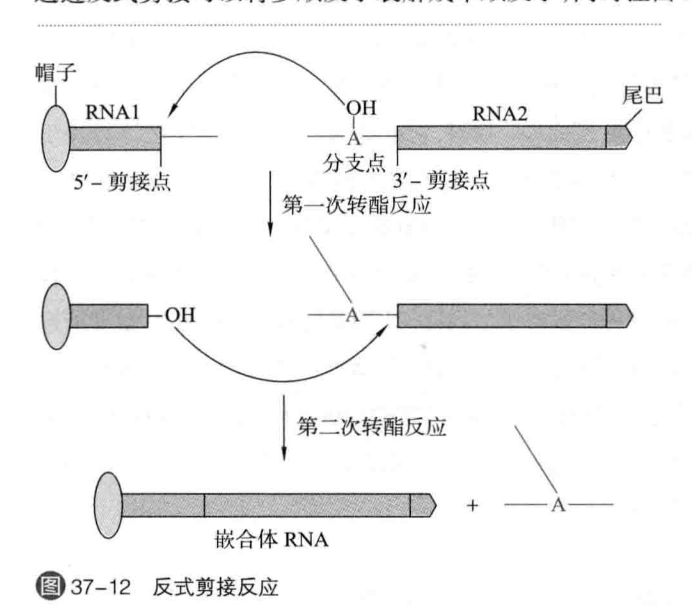


图 37－12　反式剪接反应

反式剪接也是在剪接体内发生的，即依赖于 snRNP，其本质仍然是两次转酯反应，但转酯反应是在两个不同的 RNA 分子之间发生的（图 37－12）。一般而言，进行反式剪接的 RNA 一个只有 5′－剪接点，但缺乏 3′－受体位点，另一个 RNA 有 3′－剪接点，但缺乏配套的 5′－供体位点。

2. 剪接的生物学意义

在剪接现象最初被发现的时候，科学家感到十分困惑，因为切除内含子即剪接是一个高度耗能的过程。首先在转录内含子序列的时候，就消耗了许多能量；其次在组装剪接体、进行剪接反应的时候也耗去了不少能量。因此，单从能量的角度，剪接似乎是一种浪费能量的过程，而从生物进化的角度，自然选择应该尽可能淘汰浪费的过程，除非剪接具有某种或某些特别重要的生物学功能。现在已很清楚，剪接至少具有两项重要的生物学功能：

（1）与可变剪接有关　具体就是通过可变剪接提高基因的编码能力，创造蛋白质的多样性。可变剪接可以从一个基因产生几种或多种不同的 mRNA，再经过翻译，可得到几种或多种在一级结构上有差别的多肽或蛋白质。在大多数情况下，通过可变剪接产生的蛋白质具有相似的功能，因为它们的氨基酸序列多数是相同的。但是，可变剪接也能赋予一种蛋白质具有特别的性质或功能，因为它毕竟也带来了序列的变化。例如，血管内皮生长因子（VEGF）的 mRNA 前体通过可变剪接主要产生 3 种蛋白质变体。每一种变体单独存在都能够促进血管生成，但形成的血管都有缺陷，主要表现在血管容易发生泄漏。只有 3 种变体同时存在的时候，生成的血管在形态学和功能上才是正常的。究其原因是因为 3 种变体具有互补的性质：其中最大的变体具有 2 个肝素结合位点，倾向于留在合成点附近与胞外基质结合；最小变体的 2 个肝素结合位点都丢失了，因此很容易从合成的地方扩散出去；而大小居中的变体只有 1 个肝素结合位点，这使其能以中等的速率扩散离开合成位点。综合的结果是，这三种变体同时存在能够在体内建立一种稳定的 VEGF 梯度，有利于组织新血管的健康生长。

可变剪接的优势是让一种生物用较少的基因仍然能够编码出很多不同的蛋白质。当然，不同真核生物的可变剪接的潜力是不同的，像低等的真核生物，如面包酵母，在其 6 300 个基因之中，大约只有 5% 的基因具有内含子。而含有内含子的基因也只有少数能够发生可变剪接。因此，对于单细胞真核生物而言，可变剪接并非是它们产生蛋白质多样性的主要手段。然而，对于复杂的多细胞生物来说，在很大程度上依赖于可变剪接创造蛋白质的多样性。例如，人类只有 2 万多个蛋白质的基因，但绝大多数蛋白质基因具有几个或多个内含子。保守估计，人类约 70% 的 mRNA 前体能发生可变剪接。在一些生物体内，某些基因的可变剪接“潜能”被发挥到了极限，其 mRNA 前体能产生成百上千种不同的 mRNA 变体，而如果将其他产生多样性的手段（不同启动子的选择性使用、可变加尾和编辑等）结合起来，就可进一步扩大蛋白质的多样性。例如，果蝇体内的一种叫唐氏综合征相关细胞黏着分子（Down syndrome-associate cell adhesion molecule，DSCAM）共有 111 个内含子，人体中也有类似的蛋白质。据估计，如果进行可变剪接，它就能够产生近 38 016 个 mRNA 变体，这差不多是果蝇整个基因组基因数

目的 3 倍！事实上，其中的大多数变体已被证明的确是存在的。

（2）与外显子混编（exon shuffling）有关　外显子混编这个概念最早由 Walter Gilbert 于 1977 年提出，它是指源自一个或几个基因的若干个外显子像“洗牌”一样地进行重排。基于一个外显子通常对应于一个结构域，这样很容易将不同的蛋白质分子上现成的结构域通过混编，集中到一个新的蛋白质分子上，因此这也是生物产生新基因和基因进化多样性的一种重要机制。一个典型的由外显子混编产生的新基因是在三种非洲果蝇的基因组中发现的 *jgw* 基因，它由美国芝加哥大学龙漫远教授于 2000 年发现并命名，该基因在发现的时候的年龄只有 2~5 岁。它产生的具体过程是：位于果蝇第 2 号染色体上的编码乙醇脱氢酶基因（*adh*），经反转录作用插入位于果蝇第 3 染色体的 *yande* 基因的第 3 个内含子内，从而产生 5′端由 *yande* 基因的 3 个外显子和 3′端由 *adh* 基因的 4 个外显子组成的融合基因。对 *jgw* 基因的功能研究表明，其表达产物是一种嵌合蛋白，其中 N 端来自 *yande* 基因，C 端则来自 *adh* 基因。*jgw* 基因只在雄性睾丸中表达，其蛋白质产物 JGW 可能参与雄性果蝇的激素和外激素的代谢过程。与原来的乙醇脱氢酶相比，JGW 更喜欢以长链的一级脂肪醇作为底物。龙漫远教授认为，至少有一半以上的真核生物的基因，曾经经历了类似于 *jgw* 基因的起源过程，从而证明了外显子混编以及基因重复在新基因起源过程中的普遍意义。

框 37-1　生化研究动态——U1-snRNP 的新功能

美国宾夕法尼亚大学的 Gideon Dreyfuss 等人在 2010 年 12 月 2 日的 *Nature* 上，发表了一篇题为“U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation”的论文。该论文告诉我们，参与 mRNA 剪接的 U1-snRNP 还有另外一个重要的功能，就是防止 mRNA 在转录过程中提前终止。为此，Dreyfuss 将 U1 称为真核细胞转录组的保护神，因为正是 U1 的这项功能保证了每一个转录物的完整性。

Dreyfuss 曾经发现，缺失一种帮助 snRNP 组装的 SMN 蛋白可导致 snRNP 水平和剪接发生异常。缺乏 SMN 的人会得一种常见的神经退化性疾病——脊髓性肌萎缩（spinal muscular atrophy，SMA）。

SMN 的缺乏可影响到所有的 snRNP。于是，Dreyfuss 想知道，如果细胞仅仅缺少一种 snRNP，结果会怎么样？于是他决定先研究 U1 缺失的后果。他们本来预期的结果是，细胞内没有剪接的 RNA 会增多。结果的确如此，但与此同时，他们还惊奇地发现，转录物离开起点不久就开始提前终止。

在对这些提前终止的转录物进行序列分析以后，他们发现提前终止的 mRNA 在 3′端也有多聚 A 的尾巴。要知道，多聚 A 尾巴是加尾反应（先剪切，后多聚腺苷酸化）的典型标志。这种奇怪的结果说明，缺乏 U1 造成加尾反应提前启动了。因此，Dreyfuss 认为，U1 的正常功能除了参与剪接以外，还在控制加尾反应，以保证加尾反应只有在转录到达终点以后才能进行。一个基因在正常转录的时候，U1 与移动的 RNA 聚合酶复合物一起，与新生的 RNA 转录物结合，阻止许多潜在的加尾信号被错误利用。U1 的这种功能也能解释，为什么它在细胞内的含量会比其他 snRNP 多。

（五）编辑

编辑通常是指在 mRNA 编码区内发生的任何遗传信息的改变。1986 年，Rob Benne 在研究锥体虫线粒体基因组编码的细胞色素氧化酶亚基Ⅱ（COX Ⅱ）的基因结构的时候，发现成熟的 COX Ⅱ-mRNA 在编码区比其基因的编码链序列多出 4 个 U。在排除了误测的可能性以后，Benne 认为这 4 个 U 是在转录以后通过编辑添加进去的。到了 1988 年，随着锥体虫线粒体内更多 mRNA 序列和相应的 DNA 序列被测定出，发现编辑是其线粒体基因组的一种普遍现象，且编辑的方式也不限于添加 U。

2011 年 4 月，美国宾夕法尼亚大学 Vivian Cheung 等人在进行千人基因组计划（1 000 Genomes Project）和国际单体型计划（International Hap-Map Project）中，对 27 个人的全基因组和转录组测序结果进行了分析比较，结果他们发现大量的 RNA 与其模板 DNA 序列不一致，在外显子中发现了超过 1 万个不同的位点，平均每个人大约存在 1 065 个这样的 RNA 与 DNA 序列差异。由于在不同的人中

存在相同的错配，从而表明这些序列差异并非是转录过程中的随机错误。而且他们证实这些“错配”的 RNA 也参与指导蛋白质的合成。因此，这些错配的 RNA 很可能也是通过编辑产生的，但具体的编辑机制还有待进一步研究。

1. 编辑的方式

编辑主要有两种方式：一种是在编码区内增减一定数目的核苷酸（通常是 U）；另一种是编码区的碱基发生转换或颠换，例如，C→U，A→I。

2. 编辑的机制

编辑的机制随编辑方式的不同而不同，碱基的转换或颠换需要特殊的核苷酸脱氨酶的催化，而核苷酸的插入或缺失则一般需要一种向导 RNA（guide RNA，gRNA）的介入。不管哪一种编辑，都需要一系列蛋白质和酶的参与，以识别、结合和加工编辑点，保证编辑的忠实性。

（1）依赖于特定的核苷酸脱氨酶的编辑　这里以哺乳动物小肠上皮细胞中 Apo B-48 的 mRNA 的编辑为例加以说明：哺乳动物血液里的 CM 和 LDL 分别含有两种载脂蛋白——Apo B-48（2.5×10^5）和 Apo B-100（5.1×10^5），分别在小肠上皮细胞和肝细胞内合成，而 Apo B-100 具有 Apo B-48 所具有的所有氨基酸序列。进一步研究还表明，两种蛋白质是由相同的基因编码，初级转录物也完全一样。但是，同样的初级转录物在小肠上皮细胞内经历了编辑，其 C2152 在一种胞苷酸脱氨酶催化下转变成 U，而在肝细胞中没有（图 37-13）。

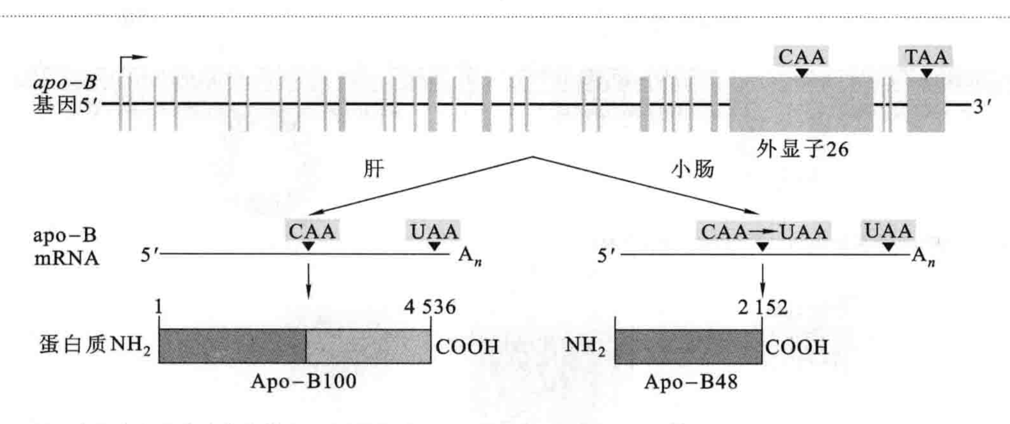


图 37-13　小肠上皮细胞内发生的 Apo B 蛋白的 Pre-mRNA 的编辑（Lodish 等，2007）

（2）依赖于 gRNA 的编辑　这种方式的编辑是在 gRNA 引导下，通过几种酶的级联作用，从 3′→5′ 方向展开。gRNA 起两个作用（图 37-14）：其一是锚定编辑点，其二是提供模板序列。在编辑过程中，GU 配对作为正常的碱基对。具体的反应为：①gRNA 与起始编辑点的周围序列配对；②特定的内切酶切开错配的碱基；③在末端尿苷酸转移酶（terminal uridylyl transferase）的催化下，切点 5′端片段的 3′-OH 插入 U，或在 3′-尿苷酸外切酶（3′-uridylate exonuclease）催化下，切点 5′端片段的 3′端缺失 U；④RNA 连接酶连接编辑点。

3. 编辑的意义

编辑的意义见图 37-15。mRNA 能通过编辑获得额外的遗传信息似乎违背“中心法则”，但归根到底还是来自 DNA。就 gRNA 指导的编辑而言，插入或缺失的序列由 gRNA 决定，而 gRNA 是从 DNA 转录来的。由核苷酸脱氨酶催化的编辑，参与的酶和蛋白质也都是由 DNA 决定的。但无论如何，编辑现象的存在时刻提醒我们，在根据基因组序列推断一种蛋白质的一级结构的时候应该谨慎行事。

当一种 mRNA 完成各种后加工反应以后，只有进入细胞质以后才能充当翻译的模板，那么细胞核内的 mRNA 是如何被运输到细胞质呢？此外，细胞又如何保证只有加工好的 mRNA（特别是剪接完成的）才能被运输出细胞核呢？

e37-1 mRNA 的运输

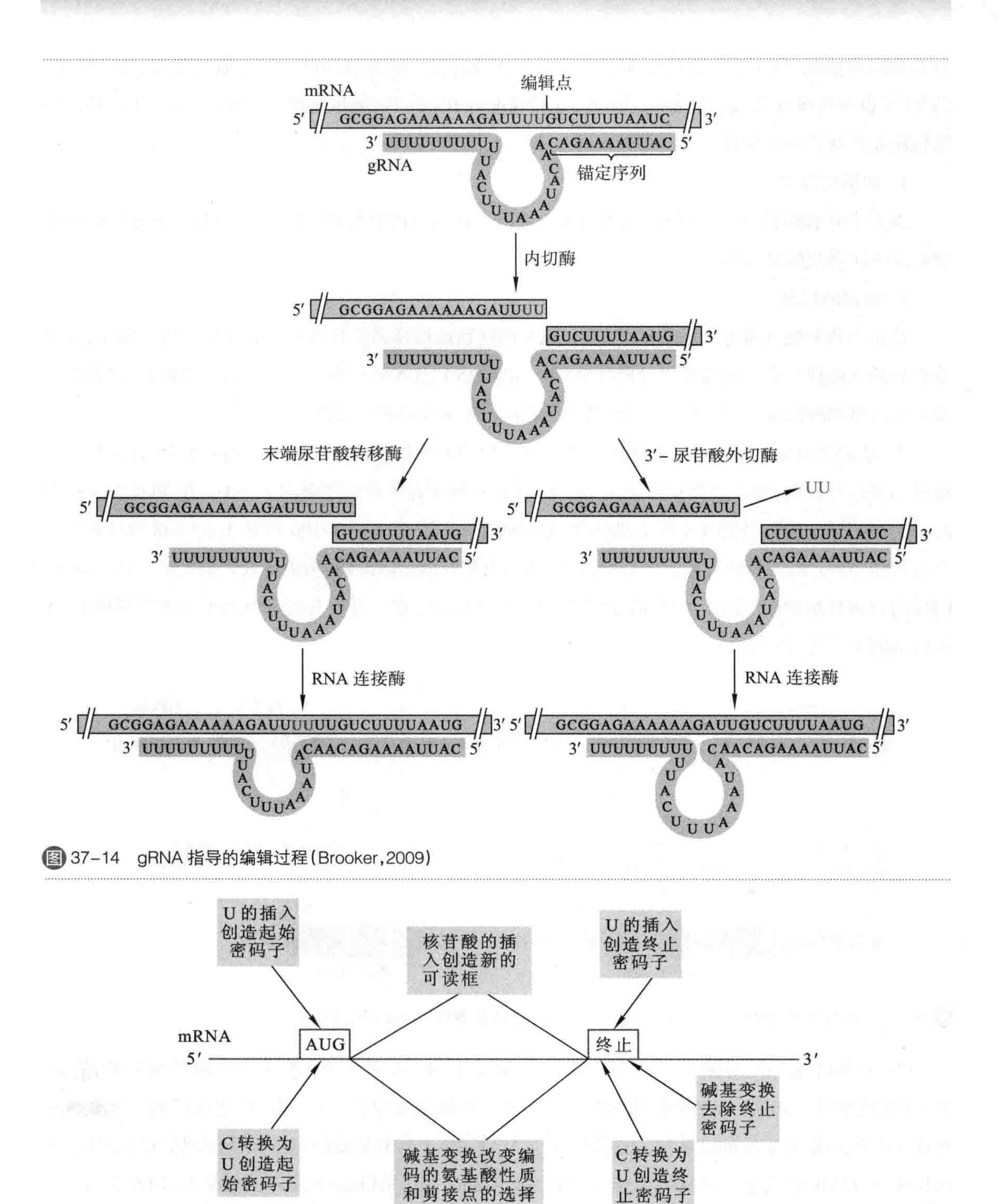


图 37-14 gRNA 指导的编辑过程（Brooker，2009）

图 37-15 编辑的意义

二、rRNA 前体的后加工

典型的真核生物的核基因组含有几百个拷贝的 rRNA 基因，它们成簇排列。其中最小的 5S rRNA 单独作为一个单顺反子由 RNA 聚合酶Ⅲ催化转录，其转录产物以 UUUUU 结尾，仅仅由 3′-外切酶做简单的修剪就可以，而 5.8S、18S 和 28S rRNA 则作为一个多顺反子，由 RNA 聚合酶Ⅰ催化转录，需要经历较为复杂的剪切和修剪，以释放出单个 rRNA。此外，成熟的 rRNA 含有大量的甲基化修饰和假尿苷 ψ，因此，真核生物 rRNA 的后加工还包括核苷酸的修饰（图 37-16）。而某些生物，如四膜虫的 rRNA 前体还含有内含子，因此这些 rRNA 前体的后加工还应包括剪接。这一类内含子被称为Ⅰ组内含子，其剪接由核酶催化，剪接的具体过程参考第十一章“核酶的结构与功能”。

真核生物 rRNA 前体的后加工需要一系列 snoRNA 的帮助。这些 snoRNA 和特定的蛋白质组装成核仁小核糖核酸蛋白颗粒（snoRNP）。某些 snoRNA 在将 rRNA 前体剪切成单个 rRNA 中起作用，但

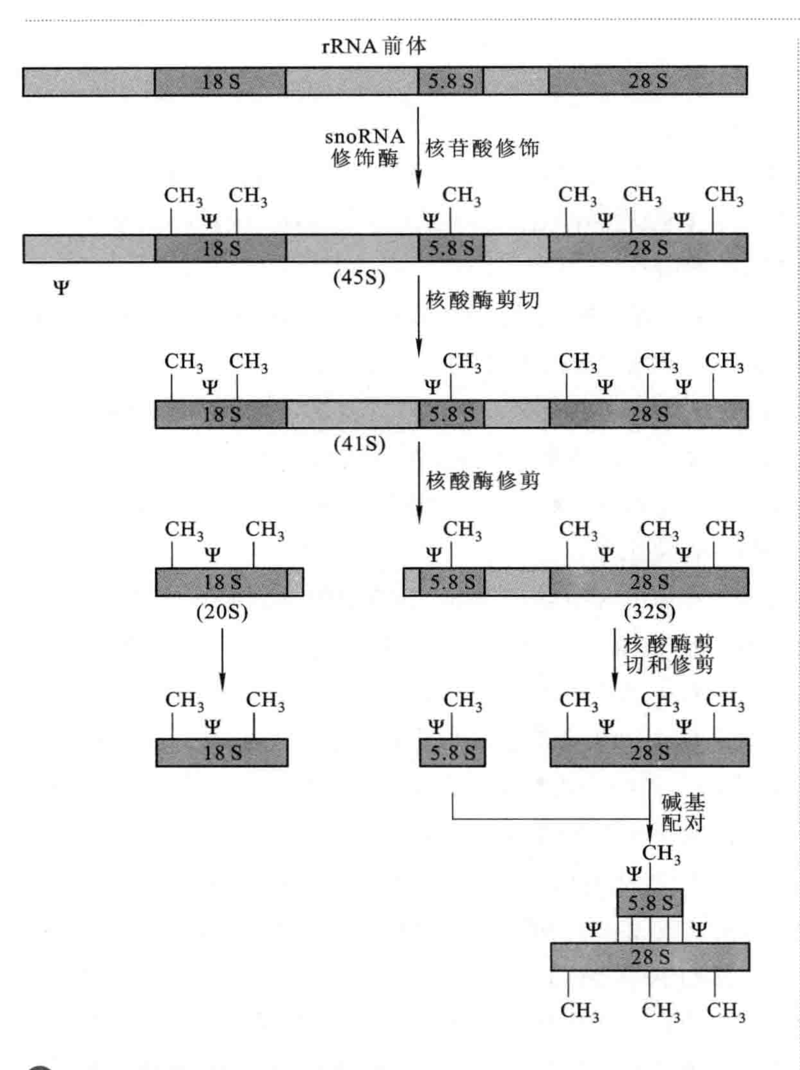


图 37-16　真核细胞核 rRNA 前体的后加工

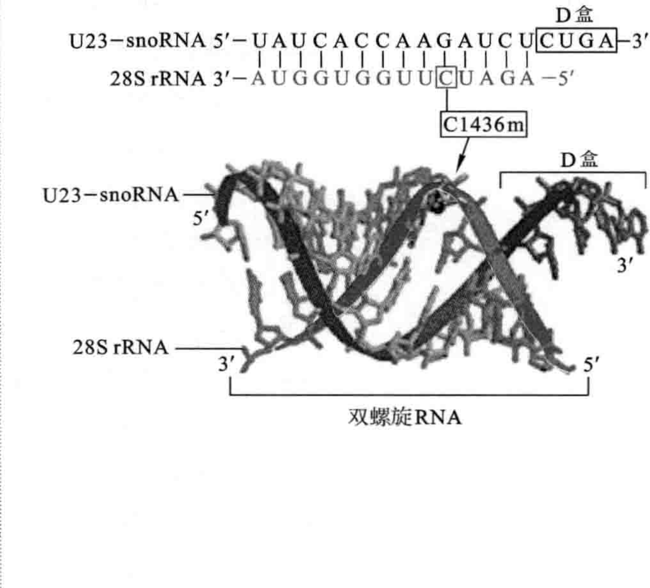


图 37-17　由 snoRNA 指导的 rRNA 前体的修饰

绝大多数 snoRNA 是通过与 rRNA 前体修饰位点周围序列的互补配对，来确定修饰位点。例如，28S rRNA 在 C1436 的核糖 2′-OH 的甲基化修饰（C1436m）是由 U23-snoRNA 决定的（图 37-17）。

一旦修饰位点确定以后，其后的修饰反应由 snoRNP 中专门的修饰酶在识别特定的 snoRNA-rRNA 复合物后进行。根据 snoRNA 上两个保守的结构元件特征，snoRNAs 一般被分为两类（图 37-18）：一类含有 C/D 盒（C/D box），它指导 2′-O 甲基化的核糖形成；另一类含有 H/ACA 盒（H/ACA box），

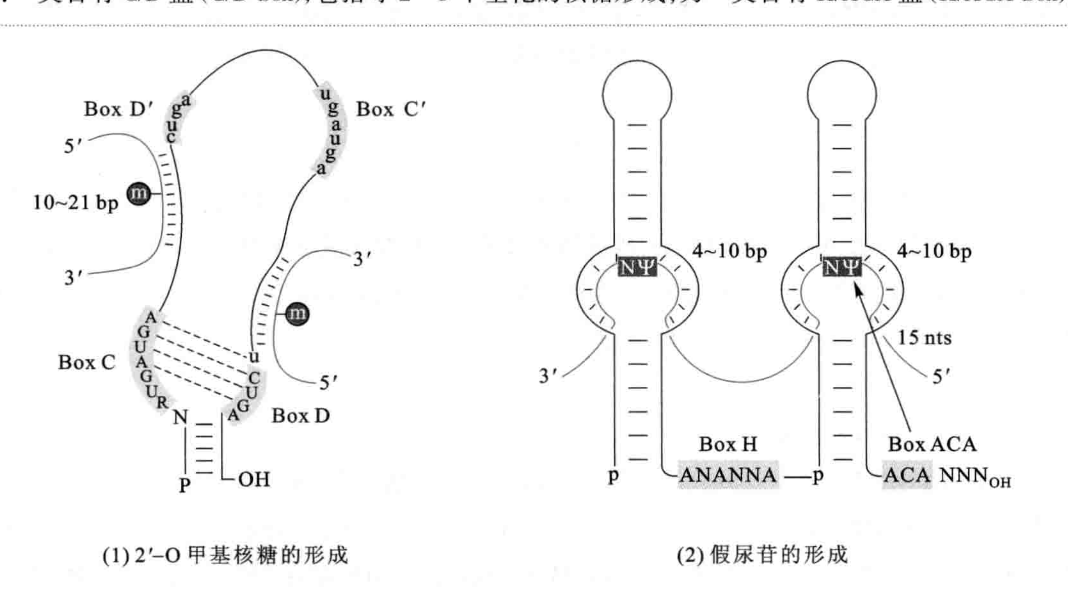


图 37-18　snoRNA 指导的 2′-O 甲基化核糖和假尿苷形成

指导假尿苷的形成。

框 37-2　生化聚焦——章鱼抗寒的秘密

看过 2010 年南非足球世界杯的人都不会忘记德国"动物预言家"——章鱼保罗(Octopus Paul),它在南非世界杯上"成功预测"了德国队所有的比赛结果,以及在最后的总决赛中西班牙的胜利,命中率达 100%,因此被网友戏称为"章鱼哥"或"章鱼帝"。

保罗生活在舒适的水族馆中,因此不必为寒冷担心。但那些生活在南极水域的章鱼呢? 显然它们是不会在触须上戴手套御寒的,但肯定在某些生理、生化机制上做出了一些适应性的改变。

最近,以波多黎各大学神经生理学家 Joshua Rosenthal 为首的研究小组就发现了其中的一种生化机制,就是章鱼可以通过 RNA 编辑,来细调神经元上钾离子通道的结构,使其能够更好地生活在低温环境中。这是科学家首次发现生物可以通过编辑,而不是通过改变特定基因的序列来适应环境的变化。该发现被在线发表在 2012 年 1 月 5 日的 *Science* 杂志上。

温度的变化可影响到神经元上离子通道的活性。当运动神经元被激活的时候,其膜上的钠离子通道会开放,致使膜电位发生去极化,从而产生动作电位。但不久后,同在膜上的一种受膜电位门控的钾离子通道会关闭,以使参与运动的神经元重新极化,为下一次神经元的再激活做好准备。这种钾离子通道的电生理性质在各种章鱼中都是高度保守的,其动力学行为表现为对温度非常敏感。其中,寒冷可延迟钾离子的关闭,从而阻止神经元再次激活。于是,包括 Rosenthal 在内的许多科学家们就猜测,生活在寒冷气候的章鱼一定会通过某种方式,改造它们的钾离子通道,以便能够更好地行使功能。

Rosenthal 本以为钾离子通道的改变源于其基因的改变。然而,当他们将生活在两种不同温度下章鱼的钾离子通道基因序列进行比对的时候,却发现生活在南极水域的章鱼和另一种生活在波多黎各水域(常年温度在 25~35℃)的章鱼,在钾离子通道的基因序列上几乎是相同的。随后,他们将两种章鱼的钾离子通道基因,导入到非洲爪蟾的卵细胞中进行表达,并测定它们在质膜上的活性。结果又发现两种钾离子通道的活性是相同的。根据这样的结果,他们预测,如果两种章鱼生活在各自的水域,在关闭钾离子通道的速率上,南极章鱼比波多黎各章鱼要慢 60 倍。

这么慢的速率会严重影响到南极章鱼的运动机能。那么南极章鱼如何能够保持它们的神经元正常工作呢? 这就是 RNA 编辑:南极章鱼能在 9 个位点进行编辑,从而在翻译水平上改变钾离子通道的氨基酸顺序。其中 I321V 位点发生的 A→I 的编辑特别重要,这种编辑可导致钾离子通道 321 位的 I 变成 V。这种变化可大大加快通道关闭的速率。研究者还测定了两种生活在热带水域的章鱼、两种温带水域的章鱼和两种南极水域的章鱼的 RNA 编辑。结果发现,生活的温度越低,钾离子通道越容易在 I321V 位点发生编辑。

这个发现的意义重大,不仅揭示了生物体增加蛋白质多样性的一种全新机制,也揭示了生物适应环境的一种新机制。

在核苷酸修饰的同时或结束以后,rRNA 前体在特定的核糖核酸酶的催化下进行剪切和修剪,其中有两种酶类似于大肠杆菌的核糖核酸酶Ⅲ和 P,还有一种核糖核酸酶 MRP。核糖核酸酶 MPR 也是一种核糖核酸蛋白颗粒,它有 2 种功能:一种是在核仁中参与 5.8S rRNA 在 5′端的后加工,另外一种功能是在线粒体 DNA 复制的时候参与切除 RNA 引物。

三、tRNA 前体的后加工

真核生物的基因组通常含有成百上千个 tRNA 的基因。这些基因的初级转录物除了在 5′端和 3′端含有多余的核苷酸序列以外,某些还具有小的内含子。内含子的位置似乎是固定的,都位于反密码子的 3′端。此外,真核生物成熟的 tRNA 也是被高度修饰的。但与细菌不同的是,真核生物 tRNA 的基因本来并没有 CCA 序列,需要通过后加工添加上去。因此,真核生物 tRNA 前体的后加工方式应包

括剪切、修剪、碱基修饰、添加 CCA 和剪接。其中剪切、修剪和碱基修饰与原核系统相似。

在真核细胞中，也发现了参与 tRNA 前体 5′端剪切的核糖核酸酶 P。但与细菌不同的是，真核细胞的核糖核酸酶 P 含有的蛋白质不止一种，例如哺乳动物有 10 种。虽然不同来源的核糖核酸酶 P 的 RNA 在一级结构上差别可能较大，但三维结构却很相似。2006 年，Ema Kikovska 等人用实验证明，与细菌和古菌一样，人的核糖核酸酶 P 也属于核酶。这意味着核糖核酸酶 P 基于 RNA 的催化在生物进化过程中得以保留。在真核生物的线粒体和叶绿体，也有核糖核酸酶 P，但人线粒体的核糖核酸酶 P 似乎缺乏 RNA 成分，这说明该核糖核酸酶 P 不可能是核酶了。

Quiz8 到此为止，你知道哪些酶同时含有 RNA 和蛋白质？这些酶中哪些属于核酶？

图 37-19 酵母 Pre-tRNA 的剪接

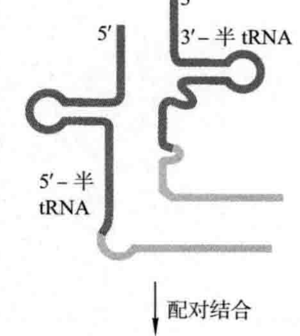
添加 CCA 是在 3′端拖尾序列被核糖核酸酶 Z 切除后进行的，由 tRNA 核苷酸转移酶催化。

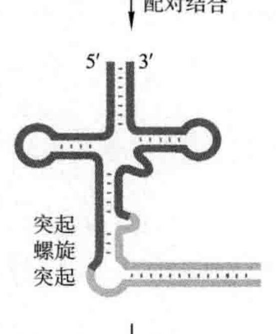
以酵母细胞的 tRNA 为例，含有内含子的 tRNA 的剪接共由 3 步反应组成(图 37-19)：①特定的 tRNA 内切酶在内含子的两端切开 tRNA，直接释放出内含子。内含子去除后在 5′-外显子的 3′端和 3′-外显子的 5′端分别留下 2′，3′-环磷酸和 5′-OH，产生 2 个半分子 tRNA。由于 tRNA 前体已形成了三叶草二级结构，因此失去内含子的 2 个半分子 tRNA 仍然通过受体茎的碱基配对结合在一起。②磷酸二酯酶切开 1 个半分子 tRNA 上的 2′，3′-环磷酸，游离出 3′-OH，激酶则在另 1 个半分子 tRNA 上的 5′-OH 上加上磷酸根。③tRNA 连接酶消耗 ATP 将 2 个半分子连接起来，磷酸酶水解 2′-磷酸根。

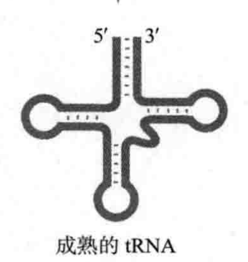
第三节 古菌的转录后加工

古菌的转录后加工在某些方面分别与细菌和真核生物一样，在某些方面仅类似于细菌或真核生物，还有少数是古菌特有的。

就 tRNA 而言，其基因的组织方式类似于细菌，即可能单独存在，也可能夹在 rRNA 基因之中。后加工的方式一般包括：①5′端和 3′端的剪切和修剪。参与 5′端和 3′端剪切的酶分别是核糖核酸酶 P 和 Z。与细菌和真核生物一样，参与 5′端剪切的核糖核酸酶 P 也属于核酶。然而，在一种叫骑行纳古菌(*Nanoarchaeum equitans*)的细胞中，没有核糖核酸酶 P。因为在这种古菌所有的 tRNA 基因的 5′端，均没有多余的前导系列，即第一个转录出来的核苷酸就是出现在成熟的 tRNA 分子中 5′端的第一个核苷酸，所以在转录以后不需要进行剪切和修剪。这也就完全摆脱了对核糖核酸酶 P 的依赖。②核苷酸的修饰，由各种修饰酶催化。③与真核生物一样，古菌绝大多数的 tRNA 基因内部无 CCA，故需要添加 CCA。④含有内含子的还需进行剪接。细菌的 tRNA 内含子属于Ⅰ组或Ⅱ组内含子，因此进行自我剪接。但古菌的 tRNA 剪接与真核生物一样，由一系列专门的蛋白质组成的酶按照一定的次序催化完成的。古菌催化剪接的内切酶识别 tRNA 分子前体上的突起 - 螺旋 - 突起的模体结构。骑行纳古菌的 $tRNA^{Glu}$ 基因由 2 个半基因组成。每个半基因含有一段内含子，其中 5′-半 tRNA 的内含子位于它的 3′端，3′-半 tRNA 的内含子位于它的 5′端。这 2 个半 tRNA 独立地转录，在转录后通过互补碱基对自发组装在一起，然后通过反式剪接合二为一(图 37-20)。







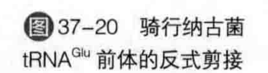


图 37-20 骑行纳古菌 $tRNA^{Glu}$ 前体的反式剪接

就 rRNA 而言，其基因以多顺反子的形式存在，通过剪切和修剪释放出 3 种 rRNA。核苷酸的修饰与真核生物一样，需要各种特殊的小 RNA（相当于真核生物的 snoRNA）与蛋白质形成的 RNP 去锁定修饰目标。但对于含有内含子的 rRNA 的剪接与其 tRNA 剪接机制相似，完全不同于真核生物。

就 mRNA 而言，其功能相关的基因一般以操纵子的形式组织在一起。其后加工类似于细菌，没有帽子结构，一般也没有多聚 A 尾巴。对于少数具有多聚 A 尾巴的 mRNA 而言，其加尾的机制和功能完全不同于真核生物，倒与含有多聚 A 尾巴的细菌相似。含有内含子的 mRNA 的剪接机制既不同于真核生物，也不同于细菌，而是类似于它的 tRNA 的剪接。

科学故事——第一例真正核酶的发现

1978 年，美国 Colorado 大学的 Tom Cech 发现，四膜虫 rRNA 的前体在成熟过程中的拼接不需要任何蛋白质的存在就可以进行。Cech 认为此催化作用来自于 rRNA 前体的本身，并将其命名为核酶。这个发现在学术界带来了极大的震撼，因为在此之前，大家都以为只有蛋白质才能充当生物催化剂。

严格地说，四膜虫的 rRNA 前体算不上真正的酶，原因是它只催化自身的反应，不像真正的酶催化其他分子的反应。另外，真正的酶在反应中自身不会被消耗，而 rRNA 前体则在反应中被水解掉。

尽管如此，Cech 的发现使许多生物学家相信，迟早有一天会发现真正的 RNA 催化剂。很快这个预言就成为了现实。1982 年，Yale 大学的 Sydney Altman 等人发现，*E. coli* 的核糖核酸酶 P 就是一种名副其实的酶。

多年以来，Sydney 一直在从事 tRNA 的结构及其前体后加工的研究。在研究 *E. coli* 的 $tRNA^{Tyr}$ 后加工反应中，他发现了参与 tRNA 前体分子 5′端后加工的核糖核酸酶 P。在电泳分离 *E. coli* 的核糖核酸酶 P 时，发现其带有高度的负电荷，当时他猜测可能带有核酸成分。几年后，他的研究生 Benjamin Stark 证实，一种大的 RNA 片段与一种大小为 1.4×10^4 的蛋白质总是同时被纯化。后来，他们将其中的 RNA 命名为 M1RNA，并发现它与 Ikemura 和 Dahlberg 在 1973 年曾报道过的一种未知功能的 RNA 是同一种东西。

为了确定 M1RNA 可能起的作用，他们首先用小球菌核酸酶处理核糖核酸酶 P，以水解其中的 RNA，结果发现酶活性完全丧失。与此同时，他们还使用蛋白酶处理核糖核酸酶 P，发现酶活性同样丧失。这似乎表明，RNA 和蛋白质都是酶活性所必需的成分。使用 CsCl 密度梯度离心，核糖核酸酶 P 的浮力密度为 1.72 g/ml，显示的是典型的 RNA 和蛋白质复合物特性。

就在 Altman 实验室在忙于分离、纯化核糖核酸酶 P 的时候，Schedl 和 Primakoff 等在对 *E. coli* 进行一系列的突变研究，结果也证明核糖核酸酶 P 中的两种成分都是体内 tRNA 前体后加工所必需的。然而到这个阶段，Altman 从来还没有想过或怀疑过，核糖核酸酶 P 中的 RNA 可能单独充当催化功能。1979 年 Ryszard Kole 发现，核糖核酸酶 P 中的 RNA 和蛋白质并不是共价结合的。当他们将两者分开来之后，都没有活性，而重新结合以后又恢复了活性。当意识到核糖核酸酶 P 的这种性质与核糖体的相似之处以后，他们开始想到，也许其中的 RNA 部分参与了酶活性中心的形成。而且，从纯化学的角度，也没有任何理由排除 RNA 就不能作为一种酶活性中心的成分或者就是一种酶的可能性。

随着 20 世纪 80 年代重组 DNA 技术的兴起，人们可以使用体外转录的方法大量制备特定的 RNA。Altman 实验室就使用了此方法大量制备了 M1RNA。与此同时，他们在纯化核糖核酸酶 P 的蛋白质亚基上也取得了很大进展。M1RNA 和蛋白质亚基的大量制备为随后的一系列实验的展开提供了便利。

1982 年夏天，Altman 与 Denver 的国家犹太医院的 Norman Pace 开始合作，共同研究核糖核酸酶 P 的结构与功能。与 Altman 不同，Pace 研究的对象是枯草杆菌的核糖核酸酶 P。他们想试试看来自不同菌种的核糖核酸酶 P 能否交换其 RNA 或蛋白质成分，而仍旧具有活性。答案是肯定的。例如，由 *E. coli* 的 RNA 亚基与枯草杆菌的蛋白质亚基重组成的杂交酶，照样可以正常地剪切 tRNA 前体。

在合作计划进行中，Pace 实验室的 Gardiner 偶然地发现，当镁离子浓度提高时（>20 mmol/L），不但正常反应增强，连对照组中不带蛋白质的 RNA 也可以单独产生催化活性。Altman 的实验室中 Cecilia Guerrier-Takada 得知此事后，也放下本来计划好的其他实验，集中全力研究这个有趣的现象。进一步研究表明，蛋白质亚基能增强酶的 k_{cat}（10~20 倍），但对 K_m 没有什么影响。由此可见，细菌核糖核酸 P 的催化作用是来自其 RNA，而不是蛋白质。接着 Altman 实验室对 M1RNA 进行了全面的研究，以确定它是否具备一种酶所应该具有的各种性质。结果也是肯定的。

Altman 的发现是对 Cech 的自我剪切 RNA 的发现的一个很好的补充，自此以后，人们意识到在生命进化史上还有一个古老而神秘的“RNA 世界”。

e37-2 本章小结填填看

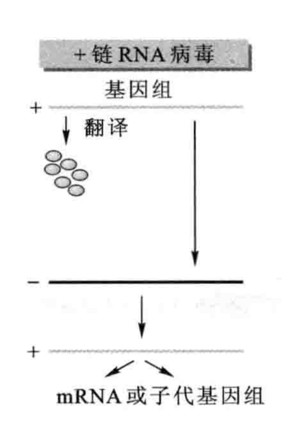
推荐网址：

1. http://en.wikipedia.org/wiki/RNA_splicing（维基百科有关 RNA 剪接的内容）
2. http://dna.kdna.ucla.edu/rna/index.aspx（美国加利福尼亚大学洛杉矶分校提供的各类 RNA 编辑信息）
3. http://bioweb.wku.edu/courses/biol22000/22mRNAprocessing/default.html（美国西肯塔基大学基础分子生物学课程，有关 mRNA 后加工的内容）
4. http://themedicalbiochemistrypage.org/rna.html#processing（完全免费的医学生物化学课程网站有关 RNA 后加工的内容）

参考文献：

1. Rosenthal J, et al. RNA editing underlies temperature adaptation in K^+ channels from Polar Octopuses. *Science*, Published Online January 5 2012.
2. Cheung V G, et al. Widespread RNA and DNA sequence differences in the human transcriptome. *Science*, DOI: 10.1126, 2011.
3. Valadkhan S, et al. Protein-free small nuclear RNAs catalyze a two-step splicing reaction. *Proc Natl Acad Sci*, 106(29): 11 901–11 906, 2009.
4. Makarov E M, et al. Small nuclear ribonucleoprotein remodeling during catalytic activation of the spliceosome. *Science*, 298: 2 205–2 208, 2002.
5. Maniatis T, Tasic B. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature*, 418: 236–243, 2002.
6. Reed R. Mechanisms of fidelity in pre-mRNA splicing. *Current Opinion in Cell Biology*, 12: 340–345, 2000.
7. Frank D N, Pace N R. Ribonuclease P: unity and diversity in a tRNA-processing ribozyme. *Annu Rev Biochem*, 67: 153–180, 1998.

第三十八章　基因组 RNA 的复制



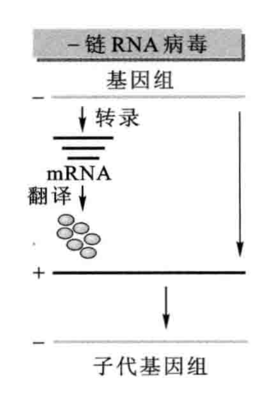


图 38-1　正链 RNA 和负链 RNA 基因组的复制

Quiz1 与 DNA 病毒相比，RNA 病毒的基因组通常很小，为 5~15 kb，你认为其中的原因是什么？

DNA 是生物主要的遗传物质，但并不是唯一的遗传物质，RNA 病毒就使用 RNA 作为遗传物质。以 RNA 为遗传物质的病毒会尽可能以最经济、最简单的方式复制自己的遗传信息，共有 2 条途径：一是依赖于 RNA 的 RNA 合成，即 RNA→RNA，另一条是以 DNA 为中间物的合成途径，即 RNA→DNA→DNA→RNA。显而易见，通过第一条途径进行复制的病毒 RNA 基因组是现在仅存的远古 RNA 世界"直系后代"。

本章主要介绍 RNA 基因组的复制，特别是逆转录病毒基因组的复制。

第一节　依赖于 RNA 的 RNA 合成

依赖于 RNA 的 RNA 合成是在依赖于 RNA 的 RNA 聚合酶（RNA-dependent RNA polymerase，RdRP）催化下进行的，该酶又称为 RNA 复制酶（RNA replicase）。RNA 指导的 RNA 合成的一般特征包括：

（1）RdRP 主要由病毒基因组编码，有的还需要宿主细胞编码的辅助蛋白　例如，Qβ 噬菌体复制酶由 4 条肽链组成，除 1 条链由自身基因组编码以外，其他 3 条链分别是宿主细胞的 S1 核糖体蛋白、翻译延伸因子 EF-Tu 和 EF-Ts。

（2）合成的方向总是 5′→3′。

（3）绝大多数在模板的一端从头启动合成，少数需要合成引物。例如脊髓灰质炎病毒（poliovirus）负链 RNA 的合成，就是以一种蛋白质分子上的一个 Tyr 残基的羟基作为引物。

（4）属于易错、高突变合成。这是因为 RdRP 无校对能力，它的错误率比 DNA 聚合酶约高 10^4 倍。如此高的错误率使得 RNA 病毒的进化速率远远高于 DNA 病毒（快 10^6 倍）！

（5）对放线菌素 D 不敏感，但对核糖核酸酶敏感。

（6）复制的场所绝大多数在宿主细胞的细胞质，少数在细胞核（例如流感病毒）。基因组 RNA 有单链和双链之别，而单链 RNA 又有正链和负链之分（以 mRNA 为标准，正链 RNA 与 mRNA 同义，负链 RNA 与 mRNA 互补），因此不同的 RNA 病毒在其基因组 RNA 复制的细节上会有所不同（图 38-1）。

一、双链 RNA 病毒的 RNA 复制

这一类病毒的基因组 RNA 为双链，如轮状病毒（rotavirus），在病毒感染到宿主细胞后，基因组 RNA 不能当 mRNA 使用。于是，病毒在包装的时候，需要将 RNA 复制酶放入到病毒颗粒之中，以便在进入宿主细胞之后，能转录出它的 mRNA。

关于双链 RNA 病毒的基因组复制，人们仍缺乏必要的了解，研究较多的是轮状病毒。这一类病毒都有双层的衣壳结构，当以内吞的方式进入宿主细胞后，部分病毒因蛋白酶的水解而脱去外层衣壳，从而在细胞质留下裸露的核心颗粒。核心颗粒中的基因组 RNA 在内部的 RdRP 的催化下，转录出具有帽子结构但无 PolyA 尾巴的单顺反子 mRNA。mRNA 在合成过程中伸入到细胞质并进行翻译。翻译出来的产物包括新的结构蛋白和 RdRP，它们以病毒质（viroplasm）的形式积累，并与全长的 mRNA 组装成非成熟的颗粒。在颗粒的内部，mRNA 作为模板，指导负链 RNA 的合成，从而得到新的基因组双链 RNA。

二、单链 RNA 病毒的 RNA 复制

(一) 正链 RNA 病毒的 RNA 复制

这一类病毒的基因组 RNA 与 mRNA 同义,因此可直接作为 mRNA 使用。一旦病毒感染进入宿主细胞,基因组就可以被翻译,如脊髓灰质炎病毒。它们的基因组 RNA 的复制由自身编码的 RNA 复制酶催化,经过互补的反基因组负链 RNA 中间物,就很容易产生新的基因组正链 RNA。

(二) 负链 RNA 病毒的 RNA 复制

这一类病毒的基因组 RNA 与 mRNA 反义,如流感病毒(influenza virus)和麻疹病毒(measles virus)。因此在进入宿主细胞之后,必须拷贝成互补的正链 RNA,才能翻译出新的病毒蛋白。于是,新病毒颗粒在形成过程中,需要将 RNA 复制酶包装到病毒颗粒之中,以便在进入新的宿主细胞之后能够转录出 mRNA。

第二节　以 DNA 为中间物的 RNA 复制

以 DNA 为中间物的 RNA 复制最为关键的一步是逆转录反应。该现象最早是在 20 世纪 60 年代末由 Howard Temin 和 David Baltimore 在研究 RNA 肿瘤病毒的复制中发现的(参看本章"科学故事")。现在已很明确,逆转录现象不仅存在于逆转录病毒中,还存在于细菌、古菌和真核生物中,其广泛性大大出乎科学家最初的预想。按照"RNA 世界"的设想,在生命从"RNA 世界"进化到"DNA 世界"中,逆转录酶可能起决定性作用。可以想象,当地球上出现了核苷酸还原酶和逆转录酶以后,从 RNA 世界进入 DNA 世界的大门就此打开了。

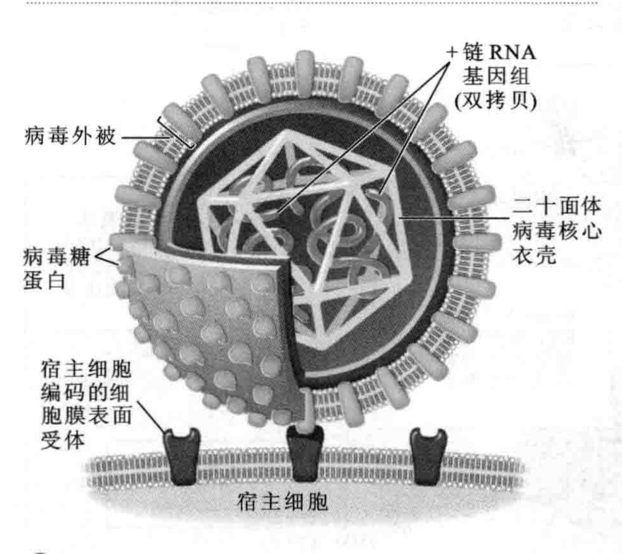


图 38-2　逆转录病毒的结构

一、逆转录病毒的 RNA 复制

逆转录病毒基因组属于正链 RNA。逆转录病毒在感染宿主细胞之后,基因组 RNA 并没有自由地释放到细胞质中,因此并不能作为 mRNA 进行翻译,而是作为逆转录酶的模板,被逆转录成 DNA。

(一) 逆转录病毒的结构

1 个典型的逆转录病毒颗粒的结构如图 38-2 所示:按照从外到内的次序,最外面的是外被(envelope),是一层衍生于宿主细胞膜的磷脂双层结构,其上含有病毒编码的表面糖蛋白(surface glycoprotein,SU)和跨膜蛋白(transmembrane protein,TM),其中 SU 为外层的外被糖蛋白,是病毒的主要抗原,TM 为成熟的外被糖蛋白的内部跨膜部分;中间是一层蛋白质核心,被称为衣壳(capsid),呈二十面体状。组成衣壳的蛋白质有基质蛋白(matrix protein,MA)、衣壳蛋白(capsid protein,CA)和核衣壳蛋白(nucleocapsid protein,NC),其中 MA 与外被的内侧面接触,CA 是病毒颗粒中含量最丰富的蛋白质,起保护核心的作用,NC 形成核心,起着保护基因组的功能;基因组 RNA 在病毒颗粒的最里面,共有 2 个拷贝,有逆转录酶、整合酶、蛋白酶和 tRNA 引物与其结合。

Quiz2 哪些生物分子的合成需要引物?哪些物质可以作为引物?

逆转录病毒的基因组 RNA 等同于一个全长的病毒 mRNA(图 38-3),其非编码序列包括 5′端的帽子结构、5′端的末端直接重复序列(repeat,R)、5′端特有的序列(5′ end unique,U5)、引物结合位点

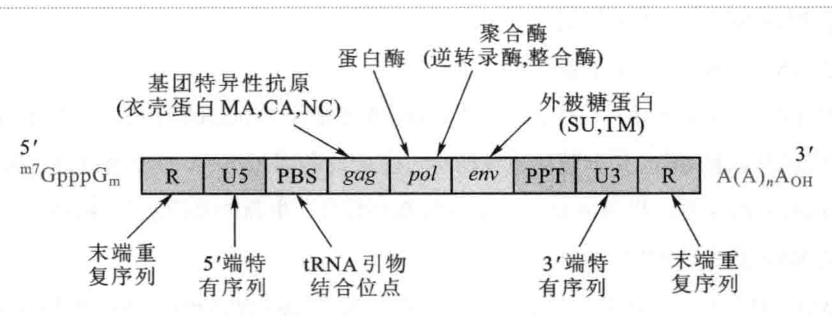


图 38-3 逆转录病毒基因组的结构

(primer-binding site,PBS)、剪接信号、引发第二条链合成的多聚嘌呤区域(polypurine tract,PPT)、3′端的多聚腺苷酸尾巴、3′端特有的序列(3′ end unique,U3)和 3′端的末端直接重复序列。编码序列通常含有 3 个结构基因,它们是编码 MA、CA 和 NC 的 *gag* 基因,编码逆转录酶、整合酶和蛋白酶的 *pol* 基因以及编码 SU 和 TM 的 *env* 基因。如果是肿瘤病毒,如劳氏肉瘤病毒(Rous sarcoma virus,RSV),就可能还含有编码癌蛋白(oncoprotein)的癌基因 *onc*。

(二) 逆转录病毒的生活史

逆转录病毒中最臭名昭著的要算 HIV(人类免疫缺陷病毒,human immunodeficiency virus)了,它是人类获得性免疫缺陷综合征(acquired immune deficiency syndrome,AIDS)即艾滋病的元凶。

HIV 可分为Ⅰ型和Ⅱ型,下面以Ⅰ型艾滋病病毒(HIV-1)为例,介绍逆转录病毒的生活史(图 38-4)。HIV-1 的生活史可分为 8 个阶段:

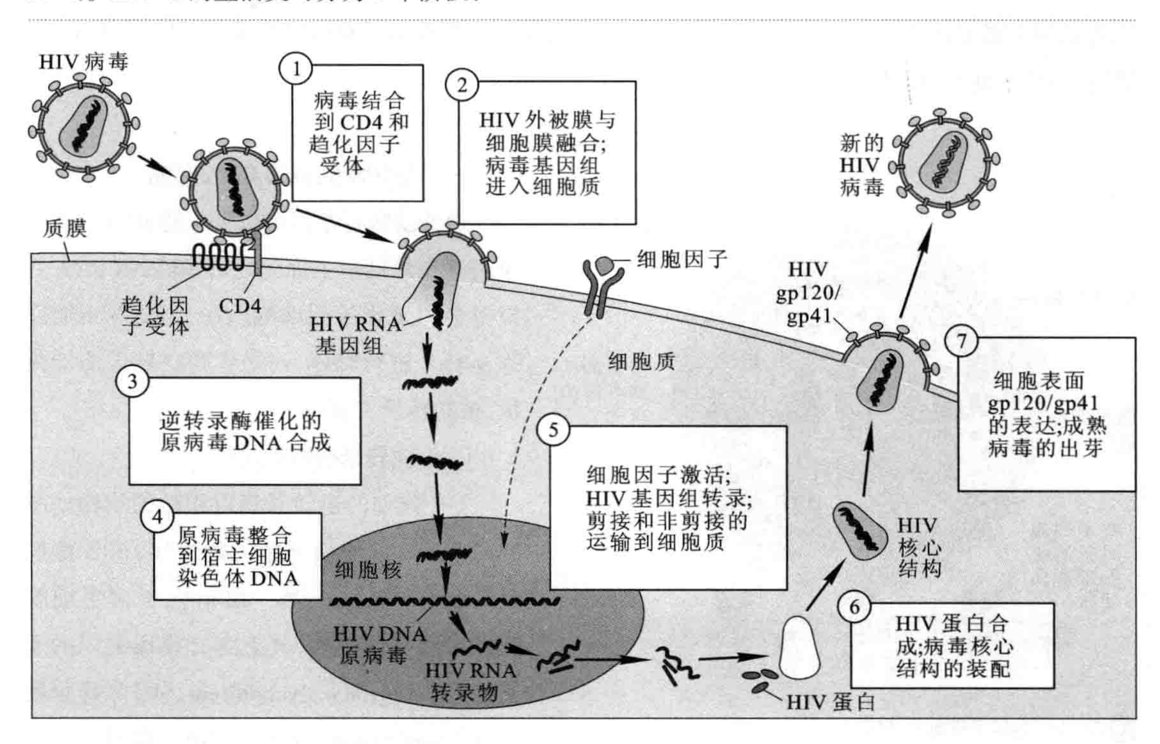


图 38-4 艾滋病病毒的生活史(Abbas,2010)

1. 病毒的附着以及和宿主细胞膜的融合

这个阶段的反应受到 HIV 在宿主细胞膜上的受体和辅助受体介导。病毒颗粒外被先特异性附着在宿主细胞膜上(图 38-5)、然后与膜融合(图 38-6)。作为受体和辅助受体的蛋白质分别是 CD4 蛋白和趋化因子(chemokine)受体——CCR5 或 CXCR4,而配体是病毒表面糖蛋白 gp120 和 gp41。其中 CD4 是一种细胞表面糖蛋白,其正常的生理功能主要是作为 T 细胞受体的辅助受体,降低细胞激活所必需的抗原数量,趋化因子的受体属于与 G 蛋白偶联的受体,其正常的功能是参与和趋化因子相关的信号传导。

HIV 的宿主细胞主要包括:CD4-T 淋巴细胞、巨噬细胞、单核细胞和树突状细胞(dendritic cell)。

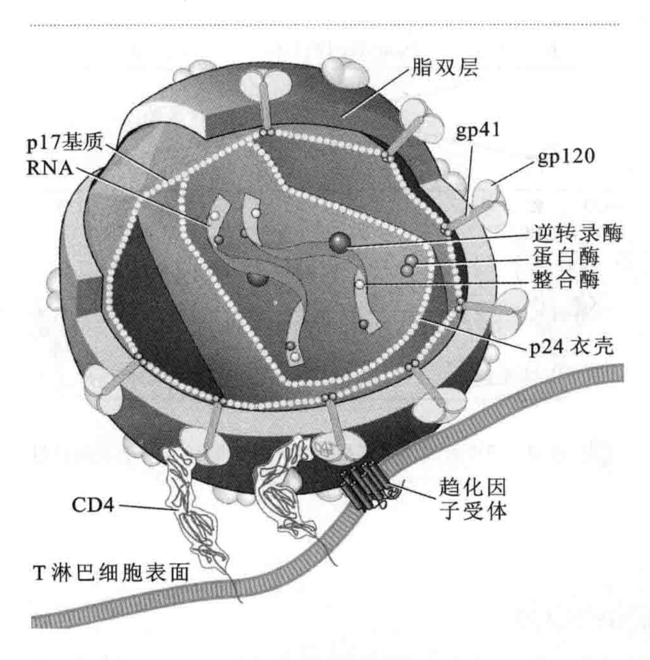


图 38-5 艾滋病病毒与宿主细胞的附着(Abbas,2010)

这些细胞都同时含有 CD4 蛋白和趋化因子受体。

2. 核心颗粒释放和逆转录

当 HIV-1 外被与宿主细胞质膜融合以后,由衣壳包被的基因组 RNA 便进入细胞质,随后在与基因组 RNA 结合的逆转录酶的催化下,进行逆转录。

逆转录酶含有大小两个亚基,小亚基是大亚基的特异性水解产物。大亚基具有 3 个酶活性:①依赖于 RNA 的 DNA 聚合酶活性,该活性用来催化负链 DNA 的合成;②核糖核酸酶 H 活性,该活性用来水解 tRNA 引物和基因组 RNA;③依赖于 DNA 的 DNA 聚合酶活性,该活性用来合成正链 DNA。小亚基的作用是保护大亚基免受水解。

Quiz3 核糖核酸酶 H 在体内和体外有哪些功能或用途?

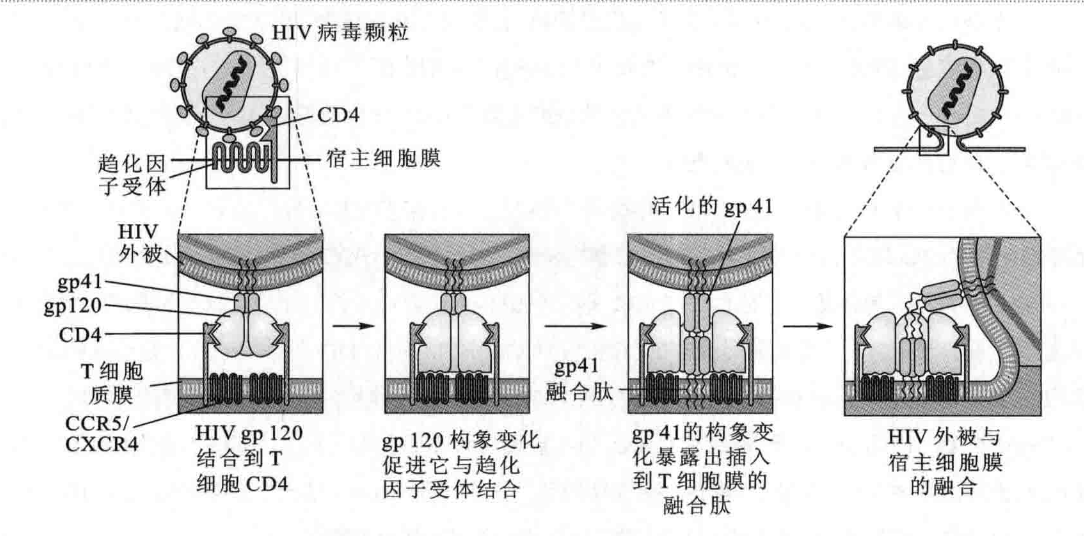


图 38-6 艾滋病病毒外被与宿主细胞膜的融合

逆转录酶催化的 DNA 合成反应的精度并不高,错误率平均为万分之一,原因是它没有 3′- 外切酶活性。这已成为 HIV 对抗艾滋病药物容易产生抗性的主要原因。

Quiz4 若是 HIV 是一种 DNA 病毒,你认为治疗艾滋病将变得更容易还是更困难?为什么?

逆转录酶是 HIV 复制必不可少的酶,而人体内除了维持染色体端粒结构完整性所必需的内源端聚酶的逆转录酶活性以外,还没有其他细胞必需的过程需要逆转录酶活性,因此,筛选到只针对 HIV 逆转录酶的强抑制剂已成为治疗艾滋病的法宝。

逆转录病毒在体内的逆转录反应共分为 7 步,通过这些反应,基因组 RNA 被逆转录成两端含有 LTR(U3-R-U5)序列的双链 DNA,原来的基因组 RNA 的 3′端加上了 U5,而 5′端加上了 U3(图 38-7)。那么为什么需要将双链 DNA 两端转变成 LTR 呢?原因在于控制逆转录病毒基因转录的启动子和增强子序列位于 U3 内部(图 38-8)。此外,只有在 5′端加上了 U3 以后,原病毒 DNA 将来才可能在宿主细胞 RNA 聚合酶Ⅱ的催化下,得到全长基因组大小的 mRNA,全长的 mRNA 再经过后加工,从而得到完整的新基因组 RNA(图 38-8)。否则,无法拷贝转录起始点上游的序列,也就得不到完整的基因组 RNA。

e38-1 逆转录病毒基因组 RNA 逆转录的详细过程

3. 原病毒 DNA 进入细胞核

原病毒 DNA 单独并不能进入细胞核,它需要先与病毒蛋白 VPR、MA 和 IN 一起组成核蛋白复合物,然后在位于 IN 上的细胞核定位信号(NLS)的指导下,通过核孔进入细胞核(参看第四十章“蛋白

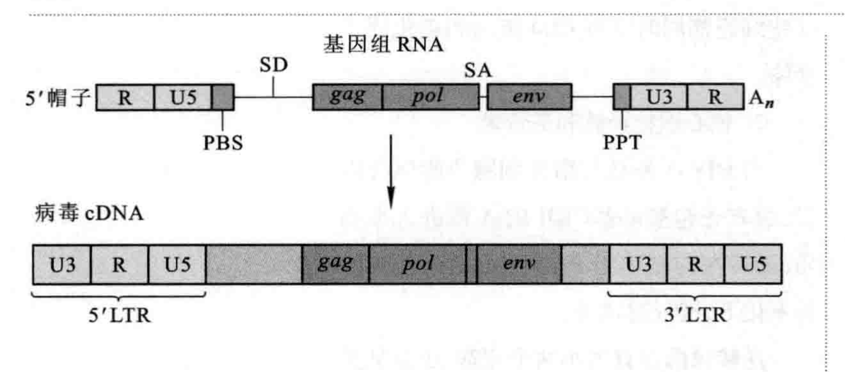


图 38-7 逆转录病毒基因组经逆转录后形成的 LTR

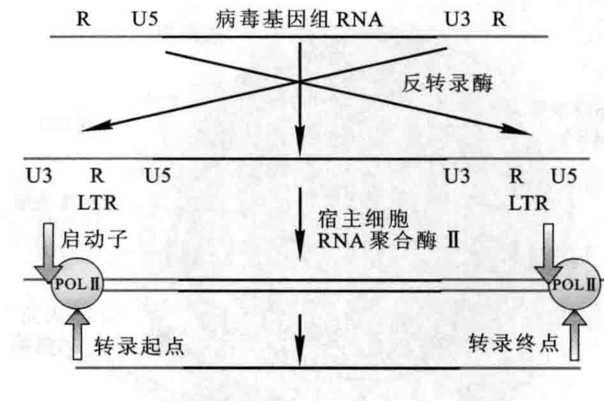


图 38-8 LTR 对维持逆转录病毒基因组 RNA 完整性的重要性

质的翻译后加工及其定向和分拣”)。

框 38-1 身边的生物化学——有天生不会感染 HIV 的人吗?

自 20 世纪 80 年代艾滋病在地球上开始泛滥以来,已有几千万人死于艾滋病。然而,并不是每一个接触 HIV 的人都会得病。科学家已经发现,人群中有少数人对 HIV 的感染似乎有抗性。这究竟是为什么呢?

答案是 HIV 就像其他病毒一样,并不能自己繁殖,必须进入宿主细胞以后才能增殖。而它进入宿主细胞又需要宿主细胞的帮助。协助 HIV 进入其内的是宿主细胞表面的 CD4 受体和 CCR5 辅助受体,这两种蛋白质缺一不可。同时带有这两种蛋白质的细胞主要是 CD4-T 淋巴细胞,HIV 感染它们以后,可逐步杀死这些细胞,而导致免疫系统的瘫痪。

许多对于 HIV 有抗性的人在 CCR5 的基因上有突变,这种突变的基因被称为 CCR5-δ32。CCR5-δ32 编码的 CCR5 较小,故不能定位到宿主细胞的表面。于是多数形式的 HIV 感染不了这样的人。一个人若是 2 个 CCR5 等位基因都是 CCR5-δ32,则几乎是不会感染 HIV 的。据估计,约 1% 来自北欧的白人就是这样的幸运者。一个人若是只有 1 个拷贝的 CCR5-δ32,则对 HIV 的感染也有一定程度的保护,即使感染了,病情也轻。这样的人约占白人人口的 20%。但黑色人种和黄色人种缺乏这样的突变。一个人即使 2 个 CCR5 等位基因都是 CCR5-δ32,也不能绝对保证不会感染 HIV。某些不常见的 HIV 可能通过其他蛋白质进入宿主细胞,尽管这种可能性很低。此外,少数 CCR5 基因完全正常的人对 HIV 也有抗性。科学家们正在对这样的人的基因组序列进行研究,以找出其中的原因。

既然 CCR5 对 HIV 的感染这么重要,并且带有缺陷 CCR5 基因的人不仅能正常生活,还对 HIV 有抗性,那么,是否可以通过阻断 CCR5 的药物来阻止 HIV 的感染呢?

答案是希望不大。原因是 HIV 是逆转录病毒,在进行复制的时候,因缺乏校对的功能很容易出差错,从而导致其基因很容易发生突变。这些突变的后代有的可能感染能力下降,有的可能感染得更好。但 CCR5 被阻断了,病毒如何进入细胞内? 答案一方面是 CCR5 被药物阻断的效率不可能 100%,这不像基因有缺陷可导致所有的 CCR5 无功能,因此,仍然有少量的 HIV 会通过没有被阻断的 CCR5 进入细胞;另一方面是一个突变的病毒可能会使用另一种辅助受体。事实上,正常的病毒感染就有这样的情况发生。在 HIV 感染的早期,病毒扩散较慢的时候使用 CCR5,而后来病毒使用另一种叫 CXCR4 的辅助受体,这时病情加重了。但病毒也许不那么容易围绕着一个已被阻断的 CCR5 进行突变,如果病毒进入不了细胞,新产生的病毒就很少,这就大大降低一种不使用 CCR5 的变体产生的可能性。

CCR5-δ32 这种天然的突变对于 HIV 感染无疑是最佳的防线,但这种突变是如何产生的呢? 另外为什么黑色人种和黄色人种几乎没有这种突变呢?

进化论告诉我们,病毒在进化,人类也在进化。但人类进化的速率显然还快不到在这么短的时间内,让很多人产生对 HIV 的抗性。换句话说,不可能是 HIV 施加的选择性压力造成这么多北欧人带有这种突

变，很可能是以前的传染病带来了这种突变的基因，正巧被用来提供对 HIV 的抗性。如果真的是这样，带有 CCR5-δ32 的人面对以前曾经爆发过的传染病更可能生存下来，并传给他们的后代。同时，具有正常 CCR5 的人群的生存率下降。在欧洲的历史上，曾有两种致命的疾病暴发过，一是黑死病（black death），二是天花，后者也需要 CCR5 和 CXCR4 才能进入细胞。已有证据表明，可能就是这两种疾病产生的选择性压力，使得一部分人获得突变的 CCR5-δ32。

4. 整合

整合即是原病毒 DNA 与宿主染色体 DNA 的共价连接，其整合位点是随机的。整合的原病毒 DNA 将成为宿主细胞永久性的遗传物质，会随着宿主 DNA 的复制而复制。

Quiz5 HIV-DNA 的整合属于位点特异性重组吗？它的整合有没有任何可能会导致细胞的癌变？为什么？

5. 转录及后加工

HIV 的转录是在宿主细胞 RNA 聚合酶Ⅱ的催化下完成的，其全长初级转录物的长度为 9 kb。转录开始于 U3-R 的连接点。然而，如果没有 Tat 蛋白，RNA 聚合酶Ⅱ的进行性就会不足，此时转录虽然仍以正常的速率进行，但将终止于 +80 nt 的位置，得到的转录物仅仅是长约 70 nt 的 TAR 序列；如果有 Tat 蛋白与 TAR 结合，周期蛋白 T 和依赖于周期蛋白的激酶 9（CDK9）就会被招募到近启动子位置，促进 CTD 的高度磷酸化，使启动子容易清空，同时增强聚合酶Ⅱ的进行性，使 RNA 聚合酶能合成出全长的初级转录物（图 38-9）。

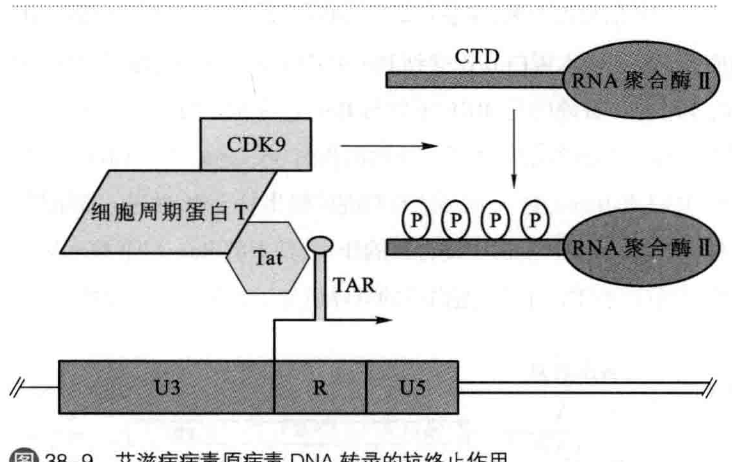


图 38-9 艾滋病病毒原病毒 DNA 转录的抗终止作用

伴随着原病毒 DNA 的转录或在转录完成以后，HIV 的转录物像宿主细胞的 mRNA 前体一样经历后加工，包括戴帽和加尾，剪接有的发生，有的不发生。如果不发生剪接，得到的就是全长的基因组 RNA，否则就是比基因组小的 RNA。全长的病毒 RNA 有 2 个功能：一是被翻译成多聚 GAG 和多聚 GAG-POL，另一个功能是被包装到子代病毒颗粒之中成为新的基因组 RNA。小于基因组大小的 mRNA 用来编码新病毒装配所需要的其他蛋白质。

简单的逆转录病毒仅有一种剪接方式，产生的 mRNA 编码 *env* 基因的产物。复杂的逆转录病毒含有多重剪接信号（图 38-10），因而会发生可变剪接，产生多种类型的 mRNA，以编码病毒特有的调节蛋白和辅助蛋白。HIV-1 的初级转录物上含有多重剪接信号，这为可变剪接提供了可能，据估计，HIV-1 通过可变剪接可以产生 30 种以上 mRNA。这些 mRNA 通过特殊的移位机制进入细胞质进行翻译。

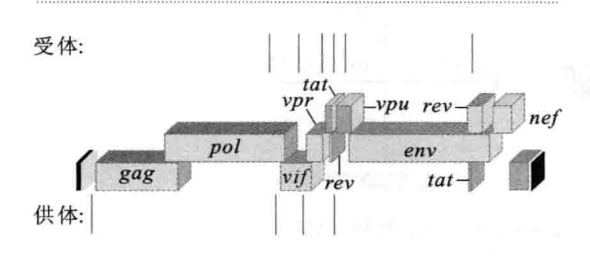


图 38-10 艾滋病病毒基因组 RNA 含有的可能剪接位点

逆转录病毒的初级转录物上还含有含有加尾信号，通常的一致序列是 AAUAAA。

由于原病毒两端 LTR 序列的存在，逆转录病毒总是含有 2 个拷贝的加尾信号，分别位于 5′端和 3′端。但是，通常只有 3′端的加尾信号才被选择使用，原因是 5′端的加尾信号受到邻近的启动子或剪接位点的抑制，而 3′端的加尾信号则受到 U3 序列的增强。

6. 转录物输出到细胞质

剪接完全的转录物很快离开细胞核，并在细胞质中进行翻译，如 *tat*、*rev* 和 *nef*，但对于没有剪接

或仅发生单一剪接反应的 mRNA 而言，它们依然含有多重剪接信号，那又如何离开细胞核呢？答案是需要一种叫 Rev 的蛋白质。Rev 与没有剪接的 RNA 上的发夹结构——Rev 应答元件（Rev Response Element，RRE）结合（图 38-11）。

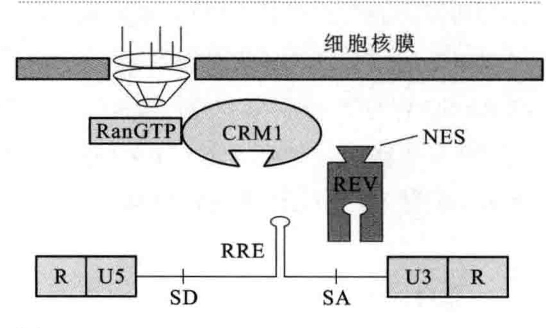


图 38-11　剪接不完全的逆转录病毒 mRNA 离开细胞核的机制

Rev 蛋白由剪接完全的病毒 mRNA 编码，因此一旦编码它的 mRNA 进入细胞质，就会很快地翻译出 Rev 蛋白。然而，Rev 蛋白具有双重信号序列：一个是富含碱性氨基酸残基的细胞核定位信号（NLS）——TRQARRNRRRRWRERQR，另一个是细胞核输出信号（nuclear export sequence，NES）——LQLPPLERLTL。那么，究竟是哪一个信号序列起作用呢？由于相对于 NES，Rev 蛋白分子上的 NLS 亲水性更强，因此它暴露得更好，故在细胞质中更容易被参与核蛋白转运的蛋白质——输入蛋白（importin）α/β 识别。于是，在细胞质中，Rev 分子上的 NLS 信号占上风，在输入蛋白 α/β 以及 Ran（一种小 G 蛋白）的帮助下，通过核孔进入细胞核。在细胞核，Rev- 输入蛋白 α/β 受到 Ran-GTP 作用而解离，随后，自由的 Rev 蛋白与含有 RRE 的病毒 mRNA 前体结合。有趣的是，RRE 正好与 Rev 上的 NLS 结合。因此，一旦 Rev 与 RNA 结合，其 NLS 信号即被屏蔽。这时 NES 便起作用了。在输出蛋白（exportin）——CRM1 的作用下，带有内含子的病毒 mRNA 与 Rev、CRM1 和 Ran-GTP 一起形成有功能的输出复合物，并进入细胞质。在进入细胞质以后，输出复合物受一种 NBP1/RanGAP 蛋白质复合物的作用，其上的 Ran-GTP 被水解成 Ran-GDP 后释放，紧接着 Rev 也被释放出来（参看第四十章“蛋白质的翻译后加工及其定向和分拣”）。

7. 翻译及翻译后加工

所有的逆转录病毒都可以产生 3 种多聚蛋白质，即多聚 GAG、多聚 GAG-POL 和多聚 ENV（图 38-12），剪接过的或非剪接过的 mRNA 从细胞核运输到细胞质后都可以作为模板进行翻译。其中 9 kb 转录物被正常翻译成多聚 GAG，或者经移框翻译出多聚 GAG-POL 融合蛋白。在病毒组装过程中，这两种多聚蛋白被多聚 GAG-POL 上的蛋白酶结构域切割成各种蛋白质单体，其中多聚 GAG 可产生 MA、CA 和 NC 等几种病毒内部结构蛋白。

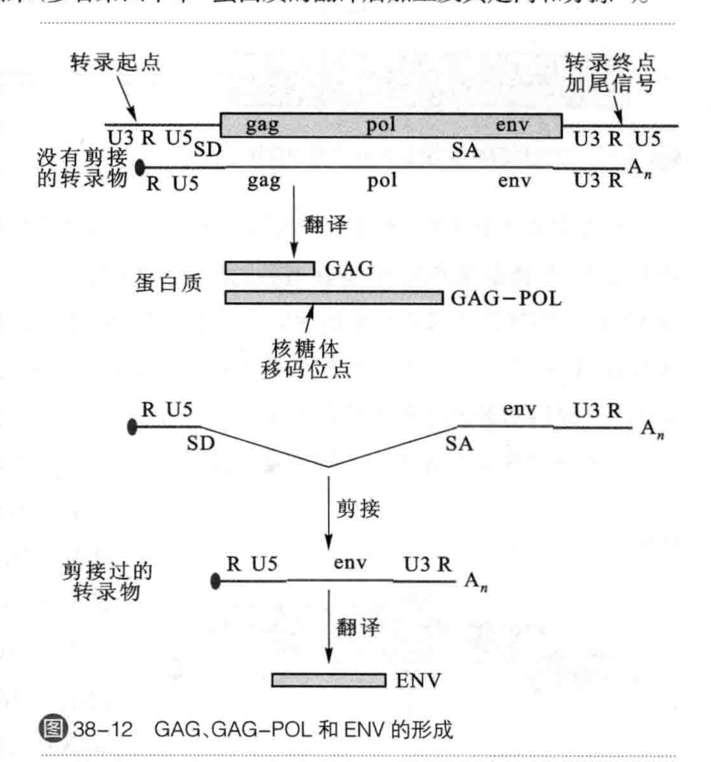


图 38-12　GAG、GAG-POL 和 ENV 的形成

一种剪接过的 mRNA（含有 *env* 基因）被翻译成 gp160 多聚蛋白。在翻译的时候，gp160 多聚蛋白进入内质网进行糖基化修饰，后被转运到高尔基体，再被宿主细胞的蛋白酶剪接成 gp120 和 gp41。

多聚 GAG-POL 除了产生 MA、CA 和 NC 以外，还能产生蛋白酶、逆转录酶和整合酶。

8. 新病毒装配和出芽释放

首先是 GAG 与 GAG-POL 的内部核心自发装配成“颗粒”；随后，2 个拷贝的 9 kb 长的基因组 RNA 随着 GAG 之中的 NC 部分，与病毒 RNA 上的包装序列结合进入颗粒之中。GAG 和 GAG-POL 上的 Myr-MA 指导复合物到达插有 gp120-gp41 的细胞质膜附近。接着颗粒与内部的 gp41 结构域结合，并开始出芽。在出芽期间或出芽以后，GAG 和 GAG-POL 被病毒的蛋白酶切割成各种单体，其

中蛋白酶、逆转录酶和整合酶与基因组 RNA 相连，其他结构蛋白构成衣壳包裹基因组 RNA。在出芽过程中，含有 gp120-gp41 的细胞质膜包被在病毒颗粒的最外面。在颗粒释放以后，GAG 和 GAG-POL 继续切割。

病毒蛋白酶虽然不是病毒装配所必需的，但却是病毒成熟和获得感染性所必需的。现在，HIV 蛋白酶的抑制剂和 HIV 逆转录酶的抑制剂(如 AZT)已成为抗艾滋病“鸡尾酒疗法”中的标准成分。

有关逆转座子的内容请参看第三十五章“DNA 重组”，有关。

Quiz6 HIV 的生活史的哪些环节可用来作为抗艾滋病药物作用的靶点？

端粒酶催化的逆转录反应的内容请参见第三十三章“DNA 复制”，在此不再赘述。

二、某些 DNA 病毒生活史中的逆转录现象

某些 DNA 病毒，例如花椰菜镶嵌病毒(Cauliflower mosaic virus，CaMV)和乙型肝炎病毒(HBV)，虽然其遗传物质是 DNA，但生活史中也有从 RNA 到 DNA 的逆转录过程。这些 DNA 病毒被称为泛逆转录病毒(pararetrovirus)。

Quiz7 写出适合 CaMV 和 HBV 的中心法则。

以 HBV 为例，其生活史如图 38-13 所示，显然它的基因组 DNA 的扩增必须通过 RNA 中间物，即必须经历逆转录。HBV 的逆转录有一个不同寻常的性质，就是其负链 DNA 的合成以 RT N 端 Tyr 残基上的羟基，而不是以核酸分子作为引物。

CaMV 的生活史与 HBV 相似，但其负链 DNA 的合成以宿主细胞的 $tRNA_i^{Met}$ 为引物。

Quiz8 哪些生物分子的合成也是以一种蛋白质分子上一个 Tyr 残基的羟基作为引物？

逆转录反应并不是真核生物及其病毒特有的，事实上，在一些原核生物体内(黏球菌、大肠杆菌和一些革兰氏阳性杆菌)也发现了逆转录现象。这些原核生物体内存在一种被称为 msDNA 的核酸。该核酸是通过特殊的逆转录反应形成的。

e38-2 msDNA 的合成

图 38-13 乙型肝炎病毒的生活史

科学故事——逆转录酶的发现

1961 年，Howard Temin 开始收集与“中心法则”不相吻合的证据。可以说，他几乎将他的全部学术生涯都花在了 RNA 肿瘤病毒的研究上。他的早期研究工作集中在劳氏肉瘤病毒(Rous sarcoma virus，RSV)上。RSV 是一种 RNA 病毒，能够将正常细胞转化成肿瘤细胞。Temin 认为，对这种病毒行为最好的解释为此病毒如何处于一种显性或前病毒状态建立一个模型。然而，既然 RNA 是出了名的不稳定，Temin 就提出了 RSV 的 RNA 基因组转变成 DNA 原病毒的模型。心里有了这个模型，他开始寻找证据去证明它。他首先得到 RSV 对 DNA 合成的抑制剂(放线菌素 D)敏感的数据，而且发现在转化的细胞里有与 RSV 基因组 RNA 互补的 DNA。对于他的模型，其他的研究者表示怀疑，看来需要一个决定性的实验来证明他的模型。

与此同时，名为 David Baltimore 的病毒学家一直在研究病毒的复制。他依靠生物化学途径，直接研究病毒本身的 RNA 和 DNA 合成。起初，他在一种非致瘤 RNA 病毒(nontumorigenic RNA virus)——疱性口炎病毒(vesicular stomatitis virus)中，分离到一种依赖于 RNA 的 RNA 聚合酶，于是他的注意力转移到 RNA 肿瘤病毒，并最终开始研究劳舍尔鼠白血病病毒(Rauscher murine leukemia virus，R-MLV)。他使用这种 RNA 肿瘤病毒，独立地证明了 Temin 的模型。

Temin 和 Baltimore 各自使用不同的途径，设计了一系列关键的实验来证明 Termin 的模型。两人都从纯的病毒样品开始，使用非离子去垢剂破裂病毒颗粒。有了破裂的病毒颗粒，他们提出了一个极为关键的问题：一种 RNA 肿瘤病毒能够进行 DNA 合成吗？为了回答这个问题，两个研究小组都加入了放射性标记的 dTTP 和其他 3 种 dNTPs（dATP、dCTP 和 dGTP）到病毒抽取物中，然后，检测有没有放射性 dTTP 参入到 DNA 分子之中。实际上，在每一个实验中，都发现有放射性标记的 dTTP 参入到核酸中。当 Baltimore 将放射性标记的 1 种 rNTP 和其他 3 种 rNTP 加到病毒抽取物时，没有检测到任何 RNA 合成。为了证明真正形成的核酸产物就是 DNA，他们分别使用专门水解 RNA 的 RNA 酶和专门水解 DNA 的 DNA 酶处理合成的产物，发现产物只对 DNA 酶敏感。这些实验结果证明病毒颗粒里含有催化 DNA 合成的酶。然而，合成 DNA 的模板是什么呢？为了一劳永逸地显示 DNA 只能从 RNA 模板合成的，Baltimore 和 Temin 都事先用 RNA 酶与病毒抽取物进行孵育，以水解可能的 RNA。如果 RNA 真是模板，那 RNA 酶的作用就导致 DNA 无法合成。事实证明，结果正是如此：RNA 酶与病毒抽取物孵育的时间越长，合成 DNA 的活性就越低。这就证明了在 RNA 肿瘤病毒里含有催化依赖于 RNA 的 DNA 聚合酶。由于这种酶催化的反应与转录相反，因此被称为逆转录酶。起初，许多科学家不愿意承认逆转录酶的存在，因为其活性违背"中心法则"，但不久对酶纯化和定性的成功让他们改变了观点。

Temin 和 Baltimor 使用不同的途径发现了逆转录酶。Temin 坚信此酶活性的存在，对于他来说，以纯化的病毒颗粒而不是感染的细胞进行生化实验，可以让他向世界证明他的模型；而 Baltimore 相信，病毒带有聚合酶的活性。他要做的事情是测定病毒有无 Temin 假定的依赖于 RNA 的 DNA 聚合酶活性。两位科学家都持有同样的信念，坚信他们的发现，而不管是不是与传统的观念相悖。

逆转录病毒的发现从许多方面影响着人类的生活。使用逆转录酶，很容易将细胞内不稳定的 mRNA 反转录成 cDNA，这就大大加速了基因的克隆和功能的研究。该发现也激励人们去寻找更多的逆转录病毒，在某种意义上，为 15 年后发现艾滋病病毒提供了方向。

Termin 和 Baltimore 的发现的重要性很快得到了世界的公认，他们因此荣获 1975 年的诺贝尔生理学或医学奖。

e38-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Reverse_transcriptase（维基百科有关逆转录酶的内容）
2. http://www.web-books.com/MoBio/Free/Ch4J1.htm（免费的分子生物学在线课程，有关 DNA 损伤修复的内容）
3. http://www.ncbi.nlm.nih.gov/books/bv.fcgi? call=bv.View..ShowTOC&rid=rv.TOC（NCBI 提供的一部关于逆转录病毒的在线教材）
4. http://en.wikipedia.org/wiki/HIV（维基百科有关 HIV 的内容）

参考文献：

1. Basavapathruni A, Anderson K S. Reverse transcription of the HIV-1 pandemic. *The FASEB Journal*, 21(14): 3 795–3 808, 2007.
2. Novembre J, et al. The evolutionary history of the CCR5-D32 HIV-resistance mutation. *Microbes and Infection*, 7: 302–309, 2005.
3. Rodriguez W V, et al. Primer-dependent synthesis by poliovirus RNA-dependent RNA polymerase. *Nucl Acids Res*, 29(13): 2 715–2 724, 2001.
4. Frankel A D, Young J A T. HIV-1: fifteen proteins and an RNA. *Annu Rev Biochem*, 67: 1–25, 1998.
5. Temin H M. The DNA provirus hypothesis: the establishment and implications of RNA-directed DNA synthesis. *Science*, 192: 1 075–1 080, 1976.

第三十九章　蛋白质的生物合成及其在细胞内的降解

蛋白质的生物合成也称为翻译，它是基因表达的最后一步。通过翻译，核酸分子中由 4 种碱基编码的遗传指令转变成蛋白质分子中由 22 种氨基酸编码的功能语言。不同的细胞对蛋白质合成的需求和能力差别很大，像哺乳动物成熟的红细胞缺乏蛋白质合成所必需的组分，因此已完全丧失了翻译的能力。但对于绝大多数细胞来说，为了生存，必须维持一定水平的蛋白质合成。整个翻译过程涉及约 300 多种不同的生物大分子，这些分子共同组装成一个高效而精确的翻译机器。

细胞在合成蛋白质的同时，也在降解蛋白质。两者处在一个动态的平衡之中。在真核细胞内，有一种依赖于泛素和蛋白酶体的蛋白质降解系统特别重要。

本章将重点介绍参与翻译过程的主要生物大分子的结构与功能、翻译的基本特征、遗传密码的性质、几大翻译系统翻译的详细过程以及蛋白质在细胞内的降解。

第一节　参与翻译的主要生物大分子

参与翻译的主要成分有核糖体、mRNA、各种氨酰 -tRNA（aminoacyl-tRNA）、一种特殊的起始 tRNA 和若干辅助性蛋白质因子（ancillary protein factor）。

一、核糖体

核糖体是一种复杂的核糖核酸蛋白质颗粒，由大、小两个亚基组成，存在于原核细胞和真核细胞的细胞质，以及真核细胞的线粒体和叶绿体两种细胞器中。各种核糖体在三维结构和功能上惊人地相似，但在具体的组分和结构上也存在着不少差别（表 39-1）。

表 39-1　核糖体的分类与组成

核糖体来源	大小	小亚基	大亚基
真核生物细胞质	80S	40S：34 种蛋白质，18S rRNA	60S：50 种蛋白质，28S，5.8S，5S rRNA
哺乳动物线粒体	55S~60S	30S~35S：12S rRNA，与大亚基共有 70~100 种蛋白质	40S~45S：16S rRNA
植物叶绿体	70S	30S：20~24 种蛋白质，16S rRNA	50S：34~38 种蛋白质，23S，5S，4.5S rRNA
细菌	70S	30S：21 种蛋白质，16S rRNA	50S：34 种蛋白质，23S，5S rRNA
古菌	70S	30S：20~30 种蛋白质，16S rRNA	50S：30~40 种蛋白质，23S，5S rRNA

（一）细菌的核糖体

细菌核糖体小亚基的大小为 30S，含有 21 种蛋白质（S1~S21）和 1 种 rRNA——16S rRNA。这种 rRNA 在 3′端含有一段富含嘧啶的序列，对于识别 mRNA 5′端的起始密码子极为重要。16S rRNA 分子上有近一半的碱基配对，大多数蛋白质正好填充在两段 RNA 螺旋之间。在与大亚基接触的表面，是一段长的无蛋白质结合的裸露 rRNA 螺旋。

Quiz1　为什么对 16S rRNA 基因序列的检测，可用于各种病原菌的检测和鉴定，以及在分子水平进行物种分类？

细菌核糖体大亚基的大小为 50S，共有 34 种蛋白质（L1~L34）和 2 种 rRNA——5S rRNA、23S rRNA。这 2 种 rRNA 都具有致密的碱基配对结构，其中 23S rRNA 具有转肽酶的活性，是一种核酶。有一种蛋白质有 4 个拷贝，其他蛋白质都是单拷贝，还有一种蛋白质（L26）与小亚基的 S20 完全一样。

（二）真核生物的核糖体

真核生物细胞质核糖体小亚基的大小为40S，含有33种蛋白质（S1~S33）和1种rRNA——18S rRNA。大亚基的大小为60S，含有50种左右的蛋白质和3种rRNA——28S、5.8S及5S rRNA。其中28S和5.8S rRNA与古菌的23S rRNA关系密切，可能起源于1个共同的远古基因，后经断裂一分为二。

真核生物的线粒体和叶绿体也含有自己的核糖体，但它们的结构与细菌更接近。

（三）古菌的核糖体

在大小和组成上，古菌的核糖体类似于细菌，含有16S、23S和5S rRNA以及50~70种蛋白质。核糖体蛋白的数目因古菌的种类不同会有所变化，进化分支早的古菌蛋白质数目多，进化分支晚的蛋白质数目少。然而，古菌的rRNA和蛋白质在一级结构上与真核生物接近。许多古菌生活在极度的环境（如极端的盐浓度和极端的温度），因此与生活在温和环境中的细菌和真核生物相比，其核糖体能更好地抵抗不良的环境。这使其特别适合做晶体研究。事实上，2009年诺贝尔化学奖得主——Ada Yonath、Thomas Steitz和Venki Ramakrishnan因对核糖体结构和功能的研究而获奖。他们的研究中使用了嗜盐古菌和嗜热古菌，并获得其核糖体高分辨率的晶体结构。核糖体重组实验显示，古菌和真核生物的细胞质核糖体在功能上可以相互取代：当将一种叫硫化裂片菌（*Sulfolobus*）的古菌核糖体大亚基与酵母的核糖体小亚基混合在一起的时候，形成的杂合核糖体在体外能进行翻译，而大肠杆菌核糖体的小亚基与这种古菌的大亚基却无法组装成有功能的核糖体。

（四）功能定位

核糖体的结构（尤其是三维结构）和功能在进化上是高度保守的（图39-1）。它之所以能够作为翻译的场所，是因为含有多个功能部位，主要包括（图39-2）：①A部位（A site）——氨酰tRNA结合部位，也称为受体部位；②P部位（P site）——肽酰tRNA及起始氨酰-tRNA的结合部位；③E部位（E site）——空载tRNA临时结合并离开核糖体的部位；④转肽酶或肽酰转移酶（peptidyl transferase）活性部位——催化肽键形成的部位，位于大亚基，由rRNA组成；⑤ mRNA结合部位；⑥多肽链离开通道（exit channel）——正在延伸的多肽链离开核糖体的通道，始于大亚基上转肽酶活性中心裂缝的底部，其直径为2 nm；⑦起始因子、延伸因子和终止因子的结合部位。

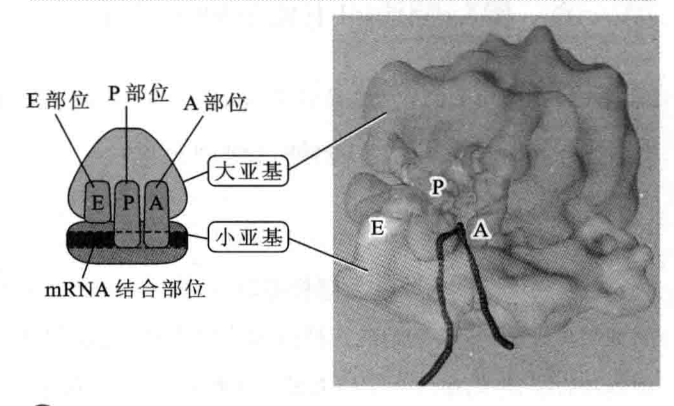


图39-1 核糖体的三维结构模型和主要的功能部位

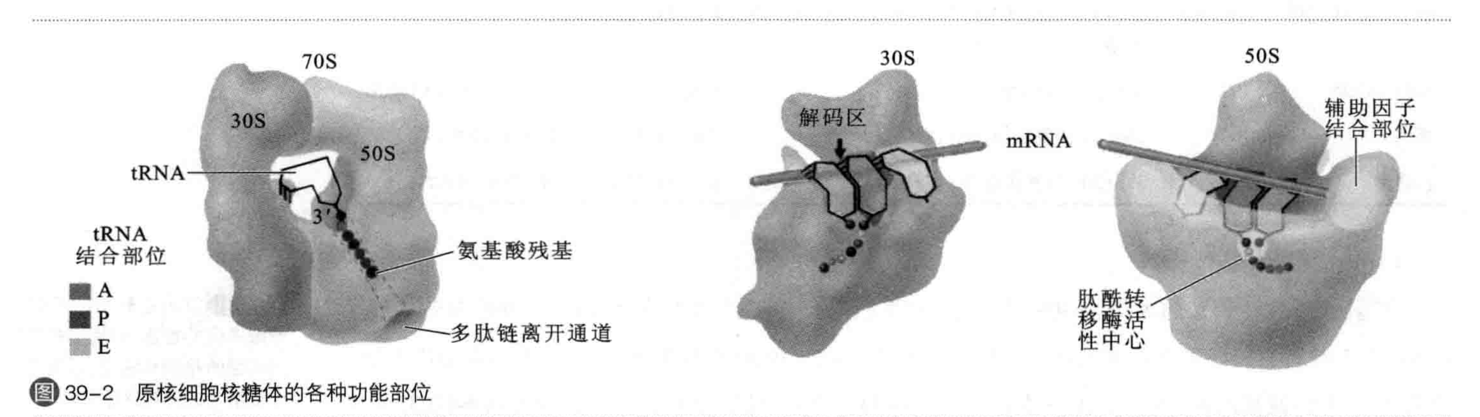


图39-2 原核细胞核糖体的各种功能部位

图39-2勾画出了一个正在翻译的70S核糖体的结构：1个tRNA分子通过一个功能端上的反密码子和mRNA上的密码子的配对，与30S亚基结合在一起，同时另一个带有氨基酸的功能端则深入到50S亚基上的转肽酶的活性中心，随时准备着作为底物参与转肽反应。新生的肽链从位于转肽酶活性中心底部的离开通道离开核糖体，这与转录过程中新生的RNA链离开RNA聚合酶的情形相似。mRNA在A

部位和 P 部位的交界处出现 1 个小小的弯曲。这个弯曲因为有 1 个金属离子的结合而得以稳定，其功能是充当一种防滑装置，有助于防止核糖体在移位的时候 mRNA 发生滑移而诱发翻译水平的移框。

图 39-2 还显示了两个核糖体亚基的结构：其上的 3 个 tRNA 结合位点显而易见，很容易想象，1 个 tRNA 分子在核糖体上从 A 部位→P 部位→E 部位的移动，以及结合在 tRNA 上的氨基酸在 A 部位，紧靠旁边与 P 部位结合的肽酰 tRNA 的肽酰基，这特别适合新肽键的形成。

（五）核糖体组装和循环

核糖体的形成是一个自组装（self-assemble）过程。先由 rRNA 和核糖体蛋白组装成大小两个亚基。在翻译的时候，再由两个亚基先后结合到 mRNA 分子上，缔合成一个完整的核糖体。这种缔合是可逆的，在翻译结束以后，核糖体又解离成大小亚基。

细胞内的核糖体还可以组装成多聚核糖体（polysome）的形式，这是指一个 mRNA 分子上同时结合几个核糖体，这种结构可以在电子显微镜下观测到（图 39-3）。形成多聚核糖体可以提高翻译的效率。而真核细胞内的核糖体还有一种与内质网结合的形式，以这种形式存在的核糖体与蛋白质翻译后的定向和分拣有关（参看第四十章“蛋白质的翻译后加工及其定向和分拣”）。

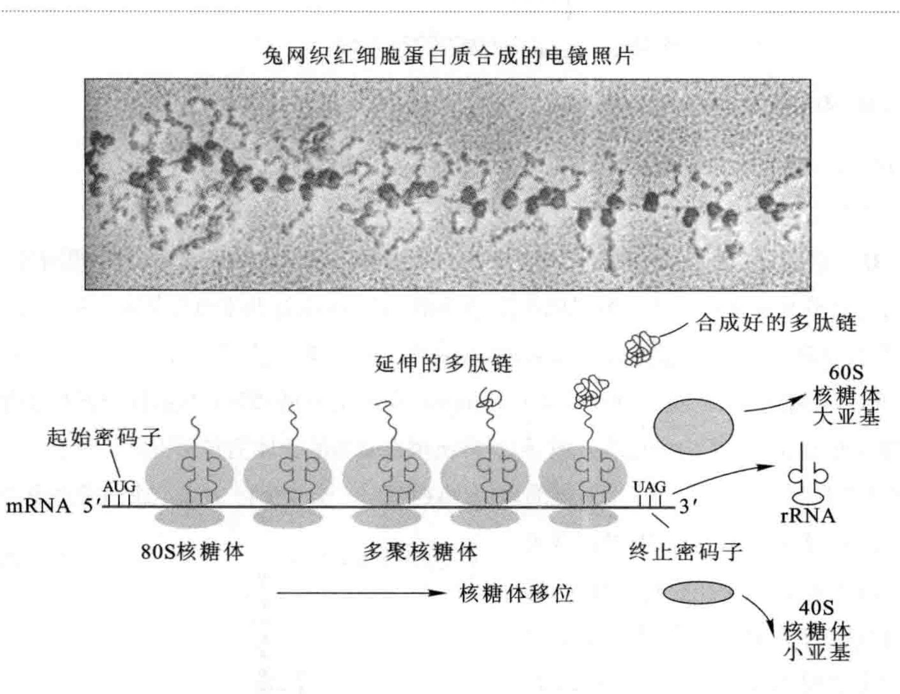


图 39-3 真核细胞多聚核糖体的结构

二、mRNA

mRNA 作为翻译的模板，至少含有一个 ORF，即以起始密码子开始、以终止密码子结束的一段连续的核苷酸序列。翻译的时候就是以这一段序列作为模板。细菌与真核生物在 mRNA 的结构上差别很大，但在两端一般都含有一段非翻译区（untranslated region，UTR）。细菌的 mRNA 通常是多顺反子——含有几个 ORF，每一个 ORF 可翻译出一种多肽或蛋白质；真核生物 mRNA 为单顺反子——只有一个 ORF，一般只翻译出一种多肽或蛋白质（图 39-4）。

三、tRNA

在翻译中，tRNA 作为一种双功能接头分子，一头与氨基酸结合，另一头通过其反密码子环上的反密码子与 mRNA 上的密码子结合，对 mRNA 进行解码。

（一）同工受体 tRNA（isoaccepting tRNA）

一个细胞中通常具有 50 多种 tRNA，负责运载 20 余种氨基酸，这意味着一种氨基酸可能对应几

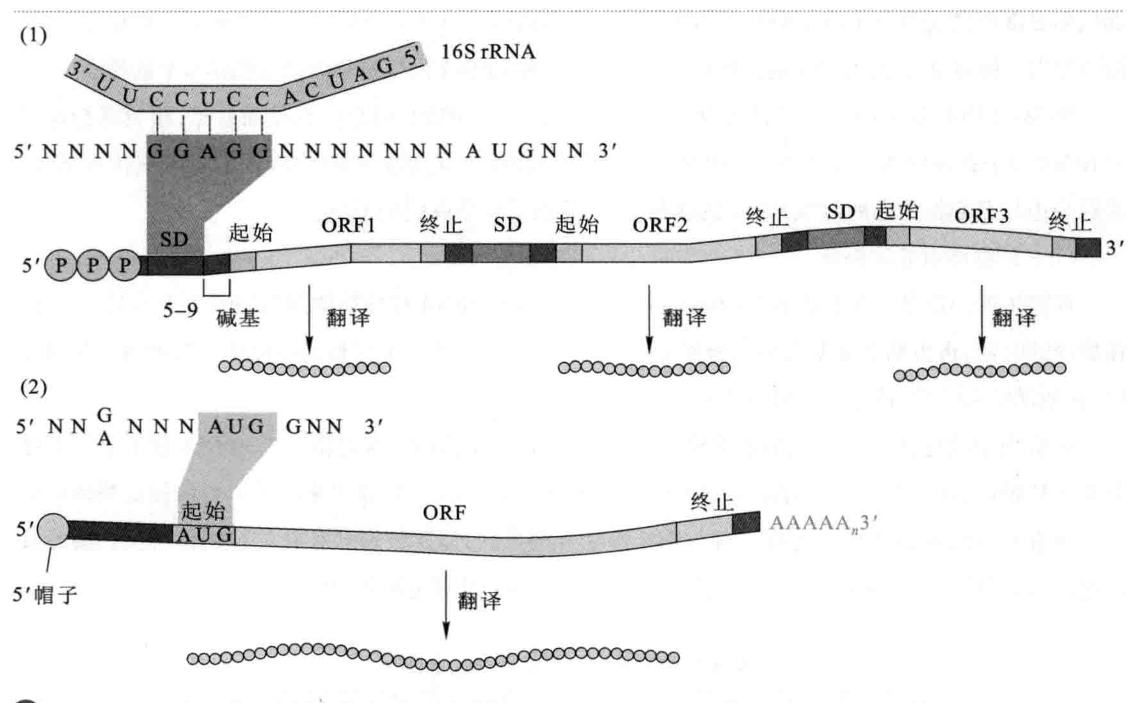


图 39-4 细菌多顺反子 mRNA 和真核生物单顺反子 mRNA 的翻译(Watson 等,2007)

种不同的 tRNA。携带同一种氨基酸的几种不同 tRNA 分子称为同工受体 tRNA。

(二) tRNA 的个性

研究 tRNA 结构与功能的关系,发现了决定某一种 tRNA 接受氨基酸专一性的主要因素,就是 tRNA 分子上由几个核苷酸甚至单个核苷酸组成的正、负元件(element),这些正负元件通常称为 tRNA 的个性(identity),有时也称为第二套遗传密码(second genetic code)。其中正元件(positive element)决定一种 tRNA 接受哪一种氨基酸,而负元件(negative element)则决定一种 tRNA 不能接受何种氨基酸。这些正负元件散布在 tRNA 分子内的很多区域,但经常出现在氨基酸受体茎和反密码子环上。

以 $tRNA^{Ala}$ 为例,其个性是一个非常规的 G3:U70 碱基对。迄今为止,所有已知序列的细胞质 $tRNA^{Ala}$ 都含有 G3:U70。若将含有 G3:C70 碱基对的 $tRNA^{Lys}$、$tRNA^{Cys}$ 和 $tRNA^{Phe}$ 突变成 G3:U70,就能使它们转而都携带 Ala;相反,若将 $tRNA^{Ala}$ 的 G3:U70 碱基对变成 G:C、A:U 或 U:G,就会使之不能再携带 Ala。有趣的是,如果能保证 G3:U70 碱基对的存在,即使仅由 24 nt 组成的 $tRNA^{Ala}$ 的微螺旋(microhelix)(图 39-5),也照样能携带 Ala。

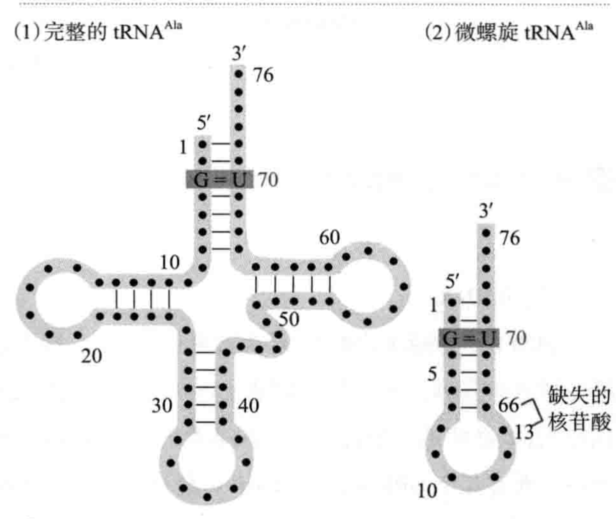


图 39-5 完整的 $tRNA^{Ala}$ 和微螺旋 $tRNA^{Ala}$ 的结构比较(Nelson 等,2008)

(三) 几种特殊的 tRNA

在 50 多种 tRNA 分子之中,有几种特殊的 tRNA:一种是专门识别起始密码子的起始 tRNA(initiator tRNA),仅参与翻译的起始。在细菌和多数线粒体内,它携带甲酰甲硫氨酸,可简写为 $tRNA_f^{Met}$,而在古菌、真核细胞和哺乳动物线粒体内则携带甲硫氨酸,可简写为 $tRNA_i^{Met}$。第二种是专门携带含硒半胱氨酸的 $tRNA^{Sec}$;第三种是在很多细菌中发现的兼有 mRNA 功能的 tmRNA。

四、氨酰 -tRNA

氨基酸在参入到多肽链之前,首先必须被活化。通过活化,氨基酸分子上的 α- 羧基连接到同源

的 tRNA 分子的 3′端，形成氨酰 -tRNA。在 tRNA 和氨基酸之间的酯键为高能键，会在后面的反应中直接用来驱动肽键的形成。活化反应由特定的氨酰 tRNA 合成酶(aminoacyl-tRNA synthetase，aaRS)催化，整个反应分为两步，与脂肪酸的活化极为相似。每活化 1 分子氨基酸，需要消耗 2 个 ATP：①氨基酸 +ATP→氨酰 -AMP+PPi；②氨酰 -AMP+tRNA→氨酰 -tRNA+AMP。产生的 PPi 会被焦磷酸酶水解，从而使得反应趋于完全。

催化反应的每一种 aaRS 面对底物都表现出高度的特异性，以确保能催化正确的 tRNA 与正确的氨基酸起反应，形成正确的氨酰 -tRNA。

Dino Moras 根据对不同的氨酰 tRNA 合成酶所进行的结构比较研究，提出将氨酰 tRNA 合成酶分成 2 类(图 39-6)：①第一类一般是单体酶，含有 1 个平行 β 折叠核心和由 2 段同源的氨基酸一致序列 HIGH 和 KMSKS 组成的 Rossman 折叠，该结构模体参与结合 ATP 和酶的催化。就此类氨酰 tRNA 合成酶而言，识别的 tRNA 个性通常包括反密码子环内的核苷酸残基和受体茎，一般在受体茎小沟一

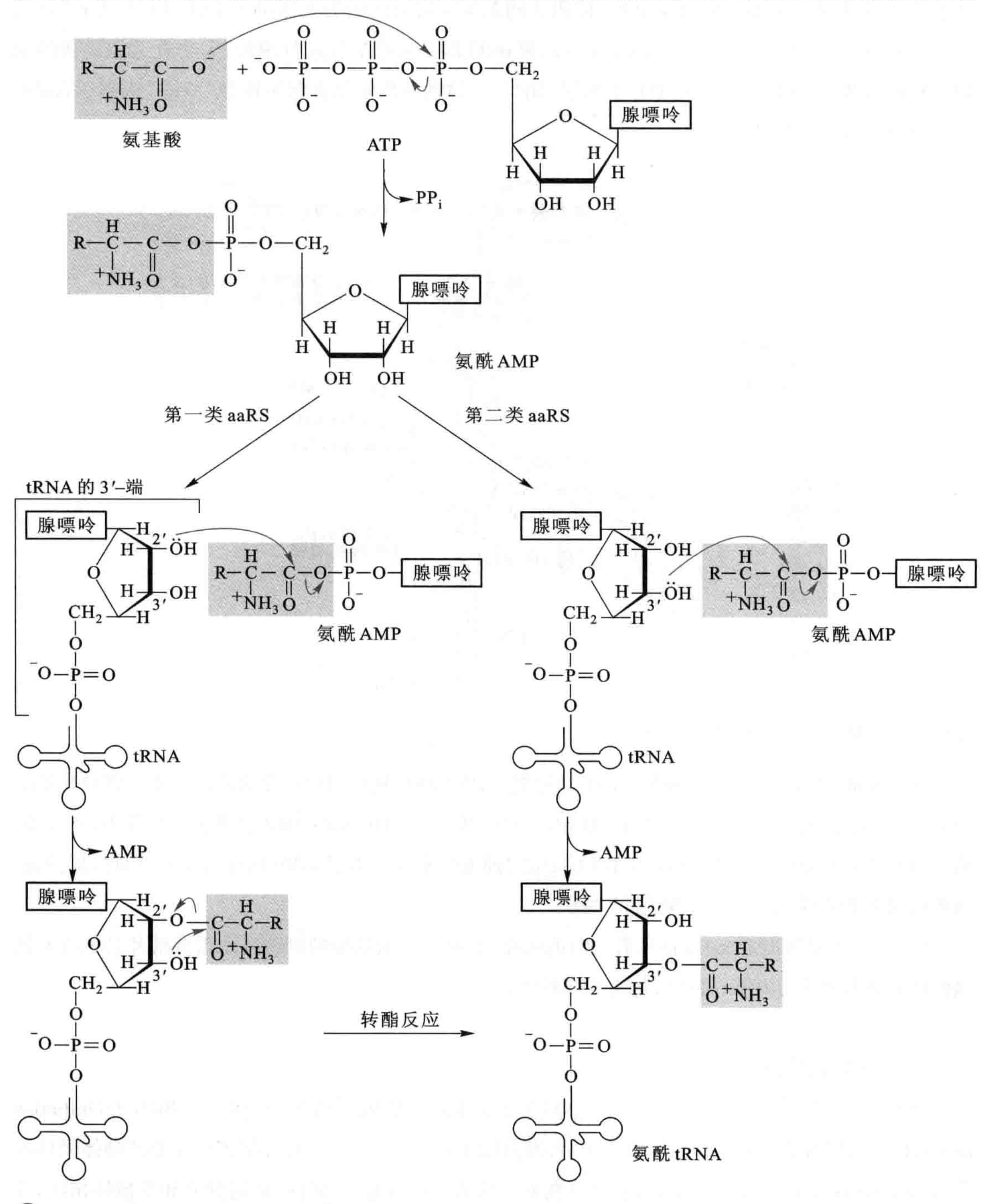


图 39-6 两类氨酰 -tRNA 合成酶的催化机理

侧与 tRNA 结合，紧握反密码子环，将 tRNA 接受氨基酸的一端置于活性中心，总是先将氨基酸转移到 tRNA 3′端腺苷酸的 2′–OH 上，然后再通过转酯反应转移到腺苷酸的 3′–OH 上。由此类酶催化的氨基酸有 Arg、Cys、Gln、Glu、Ile、Leu、Met、Trp、Tyr 和 Val。②第二类通常为寡聚酶，并没有上述氨基酸一致序列，但具有另外的保守序列，由它们形成三种首尾相连的同源结构模体，其中第一种参与二聚体的形成，另一种是"签名"模体（signature motif），由 7 段反平行的 β 折叠、3 段相邻的 α 螺旋和 1 种不多见的负责结合核苷酸的折叠组成，此模体和第三种结构模体一起组成了酶活性中心的核心部分。这类酶结合 tRNA 分子的另一面，即受体茎大沟一侧，识别的个性不包括反密码子环，最后总是将氨基酸直接转移到 tRNA3′端腺苷酸的 3′–OH 上（苯丙氨酰 –tRNA 除外）。由此类酶催化的氨基酸有 Ala、Asn、Gly、Asp、His、Lys、Phe、Pro、Ser 和 Thr。

一种 aaRS 是根据 tRNA 的个性来识别正确的 tRNA，而对正确的氨基酸的选择依赖于酶能有效识别不同氨基酸分子在 R 基团上的差异。有两种手段可用来防止生成误载的氨酰 –tRNA：一是酶活性中心优先结合正确的同源氨基酸，体积比同源氨基酸大的因为无法进入活性中心首先被排除；二是通过酶的校对中心（proof-reading site）对错误的非同源氨基酸进行编辑，大小合适的误载的氨酰 –AMP 或氨酰 –tRNA 在校对中心被水解"出局"。这两种机制结合起来称为"双筛"机制（"double sieve" mechanism）（图 39–7）。

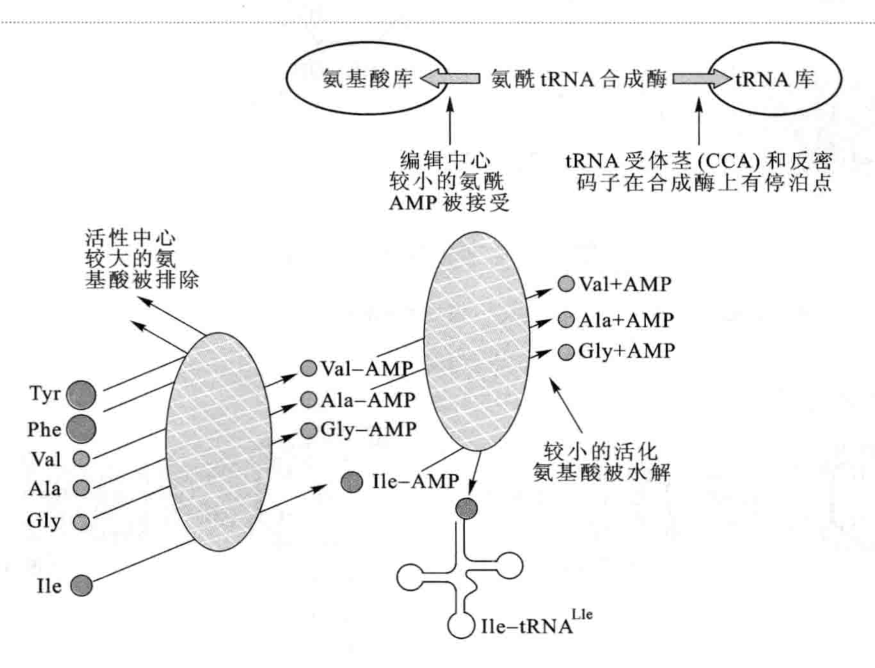


图 39–7　氨酰 –tRNA 合成酶的双筛机制

在"双筛"机制中，误载的氨酰 –AMP 或氨酰 –tRNA 转位到校对中心是必需的一步。以异亮氨酰 tRNA 合成酶为例，如果 Val 进入它的活性中心，并生成 Val–AMP，Val–AMP 会被送入校对中心进行编辑。编辑的结果是使 Val–AMP 在与 tRNA 反应之前发生水解。由于 aaRS 具有内在的校对机制，细胞内形成误载的氨酰 –tRNA 的可能性非常低。

Quiz2 你认为哪些氨基酸的活化不需要校对机制？

然而，并不是所有的 aaRS 都具有校对的功能，如果一种氨基酸的侧链基团很容易和其他所有氨基酸的侧链基团区分开来，校对机制就变得多余了。

五、辅助蛋白因子

蛋白质合成的每一步都需要一些特殊的可溶性蛋白质因子的参与，包括起始因子（initiation factor，IF）、延伸因子（elongation factor，EF）和释放因子（release factor，RF），细菌还需要核糖体循环因子（ribosome recycling factor，RRF），它们分别参与肽链合成的起始、延伸、肽链释放和核糖体循环，其中有某些辅助因子属于 G 蛋白。

第二节 翻译的一般特征

自然界存在 4 种体内翻译系统，即原核翻译系统、真核细胞质翻译系统、叶绿体翻译系统和线粒体翻译系统。所有的翻译系统都具有以下共同的特征。

1. mRNA、tRNA 和核糖体起相同的作用

所有的翻译系统都是以 mRNA 为模板，tRNA 为运输氨基酸的工具，核糖体为翻译的场所。

2. 翻译具有极性

翻译的极性表现在两个方面：一是指阅读模板时的方向性，无论是哪一种翻译系统，阅读 mRNA 的方向都是从 5′端→3′端；二是指多肽链生长的方向总是从 N 端→C 端。

Quiz3 如何设计两个实验分别证明翻译的时候，阅读 mRNA 的方向是从 5′端→3′端，而多肽链生长的方向总是从 N 端→C 端。

3. 遗传密码是三联体密码

确定遗传密码是三联体密码，即三个核苷酸编码一个氨基酸，以及三联体密码的破译曾吸引了许多生化学家的注意。他们使用了各种手段，经过不懈的努力，并最终取得了成功。

e39-1 三联体遗传密码的确定和破译

三联体密码（triplet codon）具有以下几个性质（表 39-2）：

表 39-2 标准的遗传密码表

第一个碱基	第二个碱基 U	C	A	G	第三个碱基
U	Phe	Ser	Tyr	Cys	U
	Phe	Ser	Tyr	Cys	C
	Leu	Ser	终止	终止，Sec	A
	Leu	Ser	终止，Pyl	Trp	G
C	Leu	Pro	His	Arg	U
	Leu	Pro	His	Arg	C
	Leu	Pro	Gln	Arg	A
	Leu	Pro	Gln	Arg	G
A	Ile	Thr	Asn	Ser	U
	Ile	Thr	Asn	Ser	C
	Ile	Thr	Lys	Arg	A
	Met	Thr	Lys	Arg	G
G	Val	Ala	Asp	Gly	U
	Val	Ala	Asp	Gly	C
	Val	Ala	Glu	Gly	A
	Val	Ala	Glu	Gly	G

（1）简并（degenerate）与兼职　在生物体总共 64 个密码子中，除了 3 个终止密码子 UGA、UAA 和 UAG 以外，余下的 61 个均为氨基酸密码子，编码 20 种常见的蛋白质氨基酸，这就意味着许多氨基酸的密码子不止一个。遗传密码的这种特性就叫做简并性。此外，在 61 个氨基酸密码子之中，有 3 个是兼职的，它们除了代表特定的氨基酸以外，还兼作起始密码子，这 3 个密码子是 AUG、GUG 和 UUG。其中 AUG 使用得最多；三个终止密码子有两个也是兼职的：终止密码子 UGA 和 UAG 在特定的序列环境中，可分别作为含 Sec 和 Pyl 的密码子。

Quiz4 若是使用 GUG 和 UUG 作为起始密码子，那么翻译出来的氨基酸是什么？

（2）遗传密码明确（unambiguous）而不含糊　61 个有义密码子每一个只编码一种氨基酸，没有“一码两用”或“一码多用”的情况。

（3）密码子的选定不是随机的　大多数同义密码子的差别在第三位。密码子第三个碱基的变

化倾向于编码相同的或相似的氨基酸。第二位为嘌呤的密码子大多数编码亲水氨基酸，而第二位为嘧啶的密码子大多数编码疏水氨基酸。显然密码子进化成这种样式可大大降低突变造成的危害。

(4) 通用和例外　密码子最重要的一个性质是其通用性，无论是原核生物还是真核生物都使用同一种密码子表。正因为如此，人们才能够使用原核生物来表达真核生物的基因。但是，密码子的通用性不是绝对的，有时有例外(表 39–3)，主要出现在非植物线粒体基因组，有时也出现在某些生物的细胞核基因组中。在细胞核系统中发现的例外大多与终止密码子有关，例如对于某些纤毛虫(如草履虫)，UAG 是 Gln 的密码子而不是终止密码子，迄今为止，只发现一个例外与终止密码子无关，就是一些非孢子生的酵母菌将 CUG 用做 Ser 的密码子。

表 39–3　线粒体内遗传密码的例外

生物来源	密码子				
	UGA	AUA	AGA，AGG	CUN	CCG
标准遗传密码	终止	Ile	Arg	Leu	Arg
脊椎动物	Trp	Met	终止	Leu	Arg
果蝇	Trp	Met	Ser	Leu	Arg
啤酒酵母	Trp	Met	Arg	Thr	Arg
圆酵母	Trp	Met	Arg	Thr	无
裂殖酵母	Trp	Ile	Arg	Leu	Arg
线状真菌	Trp	Ile	Arg	Leu	Arg
锥体虫	Trp	Ile	Arg	Leu	Arg
高等植物	终止	Ile	Arg	Leu	Trp

(5) 不重叠　密码子不重叠意味着在阅读同一个 ORF 时，前后两个密码子没有共用的核苷酸。

(6) 无标点　在阅读密码子的时候，从起始密码子开始，顺着 5′→3′ 的方向一个接一个阅读，中间无停顿，直至遇到终止密码子为止。

(7) 同一种氨基酸的不同密码子使用的频率不尽相同　不同的生物对同一种氨基酸的不同密码子使用的偏爱性也不同，频率出现比较低的密码子通常被称为稀有密码子。一种密码子的使用频率与其相对应的 tRNA 的丰度关系很大，一般说来使用频率越高的密码子，其对应的 tRNA 的丰度越高。

(8) 3 个终止密码子使用的频率并不相同，对紧靠终止密码子之后的第一个核苷酸的选择也不是随机的。

以大肠杆菌为例，UAA 被使用的频率最高，其次是 UAG，UGA 的使用频率最低。就位于终止密码子后的第一个核苷酸而言，细菌更喜欢使用 U，而真核生物更偏爱用 G。

4. 密码子与反密码子的相互作用决定正确的氨基酸的参入

正确的氨基酸的参入取决于 mRNA 上的密码子与 tRNA 上的反密码子之间的正确配对，与 tRNA 所携带的氨基酸无关。

5. 密码子与反密码子的相互识别遵守摆动法则

摆动法则(wobble rule)由 Crick 于 1966 年提出，用来解释一种 tRNA 反密码子如何能够识别一种氨基酸的几个同义密码子，以及某些含有稀有碱基(如次黄嘌呤)的反密码子是怎样识别由正常碱基构成的密码子的现象。该法则的内容是：密码子在反密码子配对的时候，前两对碱基严格遵守标准的碱基配对规则，第三对碱基则具有一定的自由度，具体的配对方式见表 39–4。

► 表 39-4 摆动法则

反密码子第一个碱基	密码子第三个碱基	反密码子第一个碱基	密码子第三个碱基
A	U	U	A、G
C	G	I	A、C、U
G	C、U		

以上的规则并不适用于线粒体翻译系统。在线粒体只有 21 种左右的 tRNA，如果要求它们识别 60 多种密码子，就必须遵守另外一种更为宽松的摆动法则，即超摆动法则。

摆动法则的意义在于使得翻译的过程中，tRNA 和 mRNA 更容易分离。

6. 正确的 tRNA 进入核糖体才能产生诱导契合

以细菌为例，如果正确的 tRNA 进入 A 部位，那它与 mRNA 的结合就会诱导 16SrRNA 的 A1492 和 A1493 从一段螺旋的内部环中出来，同时还造成通用保守的 G530 从原来的顺式构象变成反式构象。在新的构象之中，A1493 和 A1492 分别与由密码子 / 反密码子的前两个碱基对构成的螺旋相互作用，而 G530 同时与反密码子的第二个位置和密码子的第三个位置相互作用。这些构象诱导变化的结果可使密码子 / 反密码子的前两个碱基的配对受到核糖体的检查，以区分正确的 Watson-Crick 碱基对和错配的碱基对，而"摇摆"位置似乎能够容忍摆动法则允许的碱基对。

第三节 翻译的详细机制

蛋白质的生物合成可分为 5 个阶段，即氨基酸的活化、起始、延伸、终止和释放及折叠。也有人不把折叠与后加工归入翻译，而将它视为翻译后加工的反应。以大肠杆菌为例，翻译各个阶段所必需的成分见表 39-5：

► 表 39-5 大肠杆菌在翻译的 5 个阶段所必需的主要成分

阶段	必需成分
氨基酸的活化	20 种氨基酸，20 种氨酰 tRNA 合成酶，20 种或更多的 tRNA，ATP/Mg^{2+} mRNA，N- 甲酰甲硫氨酰 -tRNA（fMet-$tRNA_f^{Met}$），起始密码子（AUG）
起始	30S 亚基，50S 亚基，起始因子（IF1，IF2，IF3），GTP/Mg^{2+}
延伸	有功能的 70S 核糖体（起始复合物），与特定密码子对应的各种氨酰 -tRNA，延伸因子 EF-Tu、EF-Ts 和 EF-G，转肽酶，GTP/Mg^{2+}
终止与释放	终止密码子 UAG、UAA 和 UGA，释放因子 RF1、RF2 和 RF3，RRF，ATP/Mg^{2+}
多肽链的折叠	各种修饰酶，特定的剪切酶，分子伴侣等

一、细菌的蛋白质合成

细菌的翻译比真核生物的翻译要简单，以大肠杆菌为例，其各个阶段的反应如下：

(一) 氨基酸的活化

此阶段所发生的反应，主要是各种氨基酸在相应的 aaRS 催化下形成特定的氨酰 -tRNA。其中 N-甲酰甲硫氨酰 -$tRNA_f^{Met}$ 的形成需要特别注意，由两步反应组成：

1. 在甲硫氨酰 -tRNA 合成酶的催化下，形成甲硫氨酰 - 起始 tRNA

$$\text{Met} + \text{tRNA}_f^{\text{Met}} + \text{ATP} \longrightarrow \text{Met-tRNA}_f^{\text{Met}} + \text{AMP} + \text{PPi}$$

该酶也催化 Met-$tRNA_m^{Met}$ 的形成，但 Met-$tRNA_m^{Met}$ 只参与肽链的延伸，并且不会发生第二步反应。

2. 在甲酰化酶的催化下形成 N- 甲酰甲硫氨酰 -$tRNA_f^{Met}$

$$N^{10}\text{- 甲酰四氢叶酸} + \text{Met - tRNA}_f^{\text{Met}} \longrightarrow \text{四氢叶酸} + \text{fMet-tRNA}_f^{\text{Met}}$$

甲酰化反应的结果使得第一个参入的 Met 的氨基被封闭，这可能有利于保持翻译的极性，即让多肽链合成只能从 N 端向 C 端展开。

框 39-1　生化新突破——催化氨酰 -tRNA 合成的最小核酶的问世

蛋白质生物合成不可缺少的一步就是氨酰 -tRNA 的合成。在现代的细胞里，催化这一步反应的酶是蛋白质。然而在生命的进化过程中，有一个从"RNA 世界"向现代的"DNA-RNA- 蛋白质世界"的转变。在转变的过程中，最初的蛋白质合成肯定是由核酶催化的。已有足够的证据表明，现代的细胞里催化肽键形成的酶是 rRNA，而不是蛋白质。而以前也有报道，有人使用体外筛选的方法，获得了一些人工核酶，可催化氨酰基转移反应。然而，其中并没有一个核酶具有令人满意的适合在"RNA 世界"中的催化性质。这些核酶都比较大，而且都是进行自催化，而不是利用通用的水溶性底物进行反式催化，形成氨酰 -tRNA。

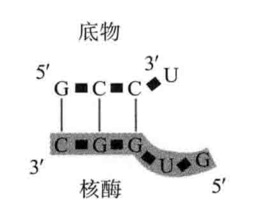


图 39-8　五核苷酸核酶与其底物的结合

不过，现在科学家终于得到了一种适合在"RNA 世界"中催化氨酰 -tRNA 合成的核酶。根据发表在 2010 年 2 月 22 日美国科学院院刊上一篇题为"Multiple translational products from a five-nucleotide ribozyme"的论文，美国的 Rebecca M. Turk 等人得到了一种目前已知最小的人工核酶，只有 5 个核苷酸组成，其活性中心只有 3 个核苷酸，但能选择性催化 Phe-AMP 中的 Phe 转移到与酶部分互补的 RNA 分子的 2′- 核糖羟基上（图 39-8）。整个反应不需要二价金属离子，对单价金属离子的依赖也有限。该核酶的发现，对"RNA 世界"假说提供了又一个重要的证据。

（二）起始阶段

这个阶段发生的主要事件是起始密码子 AUG 的识别和起始复合物的形成，共需要 3 种起始因子的帮助（表 39-6）。

表 39-6　细菌参与翻译的起始因子、延伸因子和终止因子的结构与功能

名称	功能
IF1	协助 IF3 的作用
IF2（GTP）	是一种小 G 蛋白，与 GTP 结合，促进起始 tRNA 与核糖体小亚基结合
IF3	核糖体的解离和 mRNA 的结合
EF-Tu（GTP）	是一种小 G 蛋白，其与 GTP 结合的形式促进氨酰 -tRNA 进入 A 部位
EF-Ts	是一种鸟苷酸交换因子，结合 EF-Tu，催化 GTP 取代与 EF-Tu 结合的 GDP
EF-G（GTP）	是一种小 G 蛋白，结合核糖体和 GTP，促进核糖体移位
RF1，RF2	RF1 识别 UAA 或 UAG，RF2 识别 UAA 或 UGA
RF3（GTP）	是一种小 G 蛋白，与 GTP 结合，促进 RF1 和 RF2 的作用
RRF	在翻译终止后，促进核糖体的解体

1. 起始密码子的识别

细菌翻译系统起始密码子的识别主要是依赖于 mRNA 5′端的 SD 序列（Shine-Dalgarno sequence）与 16S rRNA3′端的反 SD 序列之间的互补配对（图 39-9）。SD 序列充当细菌核糖体结合 mRNA 的起始位点，因此也被称为核糖体结合位点（ribosome-binding site，RBS）。SD 序列位于起始密码子上游约 7 bp 的区域，由 4~5 bp 组成，富含嘌呤；反 SD 富含嘧啶，两者正好可以互补配对。许多实验证明，正是 SD 序列与反 SD 序列的互补关系，才使得一个 mRNA 分子上位于 SD 序列下游的第一个 AUG 用作起始密码子。

2. 起始复合物的形成

起始复合物的形成与起始密码子的识别紧密联系在一起，起始阶段的总反应式可写成：

SD 序列　　　　起始密码子

araB	– U U U G G A U G G A G U G A A A C G A U G G C G A U U –
galE	– A G C C U A A U G G A G C G A A U U A U G A G A G U U –
lacI	– C A A U U C A G G G U G G U G A U U G U G A A A C C A –
lacZ	– U U C A C A C A G G A A A C A G C U A U G A C C A U G –
Qβ 噬菌体复制酶	– U A A C U A A G G A U G A A A U G G A U G U C U A A G –
ϕX174 噬菌体A蛋白	– A A U C U U G G A G G C U U U U U U A U G G U U C G U –
R17 噬菌体外壳蛋白	– U C A A C C G G G G U U U G A A G C A U G G C U U C U –
核糖体蛋白S12	– A A A A C C A G G A G C U A U U U A A U G G C A A G A –
核糖体蛋白L10	– C U A C C A G G A G C A A A G C U A A U G G C U U U A –
trpE	– C A A A A U U A G A G A A U A A C A A U G C A A A C A –
trpL 前导肽	– G U A A A A A G G G U A U C G A C A A U G A A A G C A –
16 S rRNA 3′-端	3′ H O A U U C C U C C A C U A G – 5′

图 39-9　SD 序列和反 SD 序列（Campbell 等，2009）

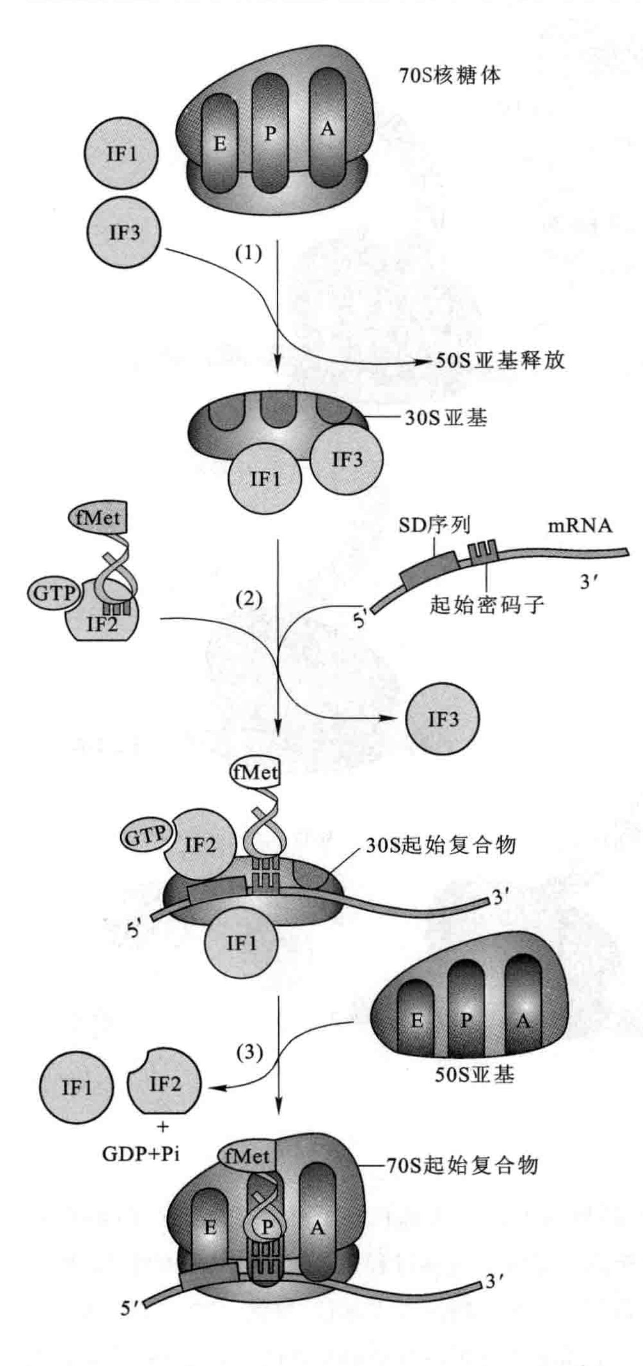


图 39-10　细菌蛋白质合成的起始

$$30S + 50S + mRNA + fMet\text{-}tRNA_f^{Met} + IF1 + IF2 + IF3 + GTP \longrightarrow [70S \cdot mRNA \cdot fMet\text{-}tRNA_f^{Met}] + GDP + Pi + IF1 + IF2 + IF3$$

在形成的三元复合物中，mRNA 像一根细线一样，穿过小亚基上弯曲的通道，其可读框内的序列将作为模板指导翻译。其中起始密码子正好处于 P 部位的底部，而 fMet-tRNA$_f^{Met}$ 通过密码子与反密码子的互补作用定位于 P 部位，第二个密码子落在 A 部位的底部，随时准备与进入 A 部位的一个氨酰-tRNA 分子上的反密码子相互作用。

形成三元起始复合物的具体过程如图 39-10 所示：首先，在 IF1 的帮助下，IF3 与 30S 亚基结合，导致大小亚基解离；随后，mRNA 模板在 IF3 的帮助下与核糖体结合，并通过 SD 序列与 16S rRNA 在 3′端的反 SD 序列之间的互补配对，使得起始密码子处在合适的位置。而 fMet-tRNA$_f^{Met}$ 则在 IF2·GTP 的帮助下与小亚基结合，但 fMet-tRNA$_f^{Met}$ 并不是通过密码子与反密码子的互补配对才结合上去的，而是因为 IF1 的结合阻断了小亚基上的 A 部位结构域，致使 fMet-tRNA$_f^{Met}$ 只能结合到 P 部位结构域上。在 fMet-tRNA$_f^{Met}$·IF2·GTP 结合上去以后，50S 亚基随后加入，并作为 GAP 激活 IF2 所具有的 GTP 酶活性。一旦 IF2 的 GTP 酶受到激活，与它结合的 GTP 随之水解成 GDP 和 Pi。GTP 的水解是一个重要的信号，它直接诱发 3 种起始因子的释放。

Quiz5 如果使用不能水解的 GTP 类似物代替 GTP 与 IF2 结合，你认为将对翻译的起始有何影响？

一旦起始因子释放，有功能的三元起始复合物便形成了。

在形成起始复合物的过程中，有两点值得注意：一是 GTP 水解的意义。从上面的叙述中可以看出，起始 tRNA 与核糖体的结合并不消耗能量，水解 GTP 使得起始因子的释放得以发生。这为延伸因子的结合铺平了道路。二是 fMet-$tRNA_f^{Met}$ 在起始复合物中的位置。它是唯一直接进入 P 部位的氨酰 -tRNA，其他种类的氨酰 -tRNA 都是先进入 A 部位。

（三）延伸阶段

在起始复合物形成以后，翻译即进入延伸阶段，延伸阶段所发生的主要事件是进位、转肽和移位且不断的循环（图 39-11）。

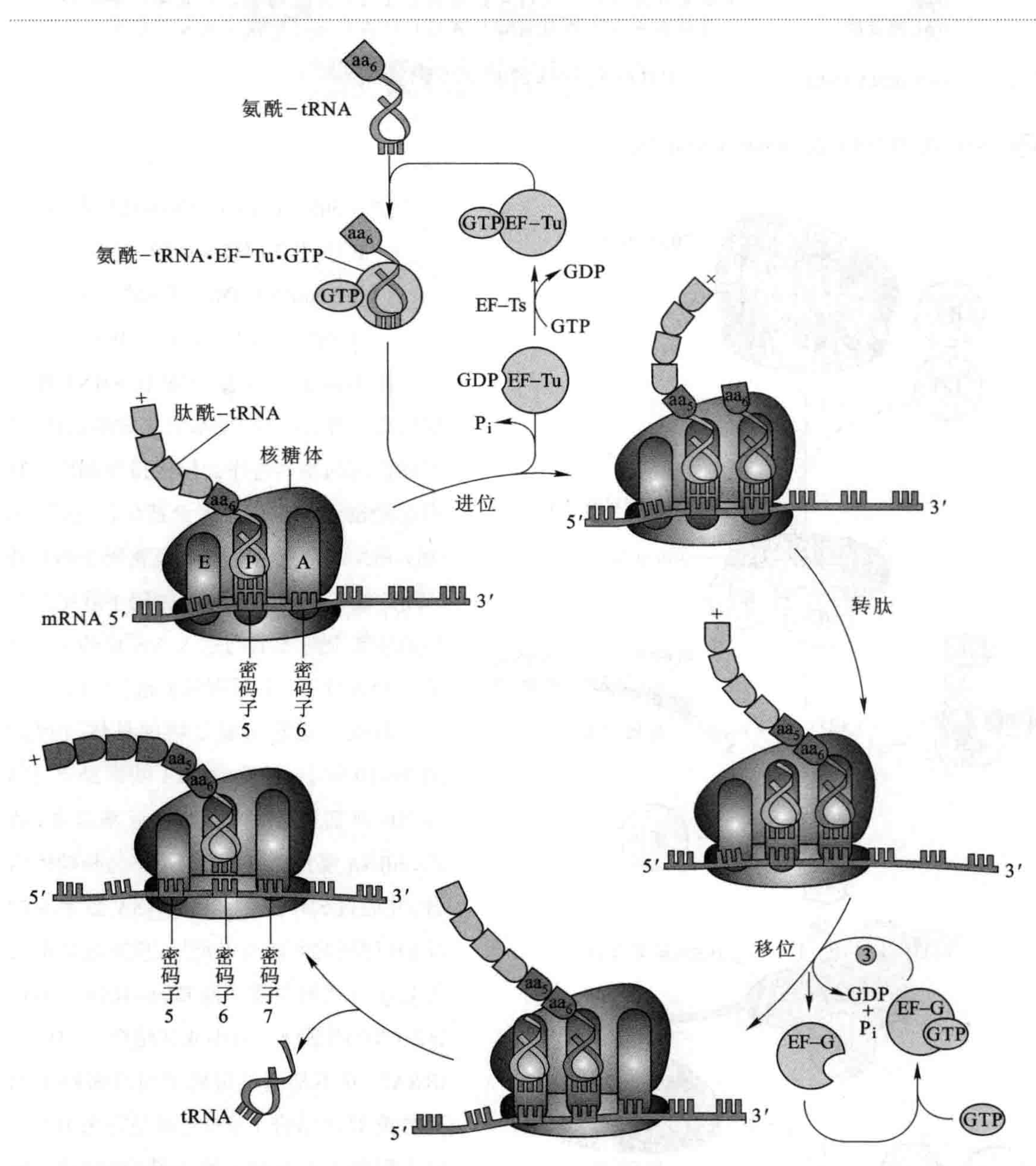


图 39-11 细菌翻译延伸阶段的反应

1. 进位（entry）

是指正确的氨酰 -tRNA 在 EF-Tu·GTP 的帮助下进入 A 部位。虽然在没有 EF-Tu 的存在下，氨酰 -tRNA 也能进入 A 部位，但是效率极低。进位的具体过程是：首先 EF-Tu·GTP 和 fMet-$tRNA_f^{Met}$ 以外的氨酰 -tRNA 形成三元复合物，随后三元复合物进入 A 部位；氨酰 -tRNA 的结合会诱导小亚基的构象发生变化，致使 16S rRNA 上一段高度保守的碱基能够与密码子 / 反密码子复合物前两个碱基对形成的双螺旋的小沟发生“亲密”接触，而有利于确保正确的 tRNA 的结合。如果上述

相互作用无法稳定特定的核糖体构象，就意味着错误的氨酰 -tRNA 进入了 A 部位，它就会在肽键形成之前被释放。

核糖体上的一个结构域充当 EF-Tu 的 GAP，而 GAP 能否及时被激活依赖于密码子 / 反密码子的相互识别。一旦正确的氨酰 -tRNA 进入 A 部位，EF-Tu 所具有的 GTP 酶活性不久即被激活，与它结合的 GTP 水解成 GDP 和 Pi。当 GTP 水解以后，EF-Tu 的构象发生较大的变化，这种构象的变化促进了 EF-Tu 的释放。而 EF-Tu 的释放引起氨酰 -tRNA 在核糖体上重新排布，为进入转肽反应创造条件。

释放出来的 EF-Tu·GDP 在 EF-Ts 的催化下，由细胞质基质的 GTP 取代 GDP 而重新转变成为有活性的 EF-Tu·GTP。

与起始阶段相似，氨酰 -tRNA 进入 A 部位并不需要 GTP 的水解。有人使用不能水解的 GTP 类似物[如 GDP(NP)]代替 GTP，发现氨酰 -tRNA 照样能够进入 A 部位。在这里，GTP 的水解同样具有特殊的意义——它引发了 EF-Tu 与核糖体的解离。实际上，EF-Tu 通过两种方式来保证正确的氨酰 -tRNA 能各就各位：①阻止氨酰 -tRNA 的氨基端进入核糖体的 A 部位（被携带的氨基酸并不参与校对）。这确保了氨酰 -tRNA 上的反密码子和 mRNA 上的密码子之间配对在先，防止误载的氨酰 -tRNA 与 A 部位不可逆地结合以及随后形成错误的肽键。②GTP 缓慢水解，只有在 GTP 水解以后，EF-Tu 才会离开核糖体。在与 EF-Tu 结合的 GTP 尚没有水解之前的一段时间内，反密码子与密码子之间进行相互校对。因此，EF-Tu 的 GTP 酶活性越低，密码子与反密码子相互校对的时间就越长，错误的氨酰 -tRNA 被 A 部位容纳的可能性就越小，肽链延伸的速率也会降低。于是，EF-Tu 的 GTP 酶活性相当于一个计时器，由它决定翻译的速率与忠实性的平衡。

2. 转肽

当正确的氨酰 -tRNA 进入 A 部位以后，紧接着就发生由转肽酶催化的转肽（transpeptidation）反应——与 A 部位结合的氨酰 -tRNA 上的氨基 N 去亲核进攻结合在 P 部位上的肽酰基或氨酰基（位于肽酰 -tRNA 或氨酰 -tRNA 上），并形成肽键（图 39-12）。如果是第一个肽键，那一定是在甲酰甲硫氨酸和第二个氨基酸之间形成的。

转肽酶是一种核酶。支持这种观点的主要证据有：①在核糖体上还没有发现一种蛋白质单独或者和其他蛋白质一起催化肽键的形成。②小亚基 16S rRNA 特殊的区域在 A 部位和 P 部位与反密码

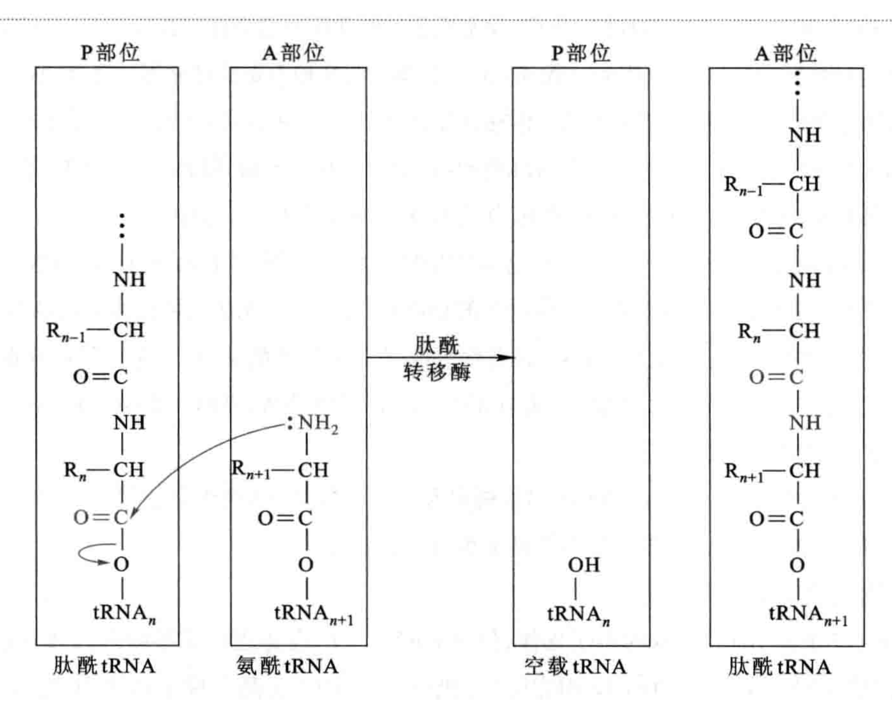


图 39-12 蛋白质合成过程中的转肽反应

子区域相互作用。相反，大亚基 23S rRNA 与肽酰 -tRNA 的 3′-CCA 相互作用，从而使之位于转肽酶合适的位置。③红霉素和氯霉素抑制转肽酶的活性，但某些 23S rRNA 序列发生突变的菌株对这两种抗生素具有抗性。④核糖体的三维结构显示转肽酶的活性中心由 RNA 组成，最近的蛋白质离活性中心有 2 nm，这样的距离太远，不可能参与催化。⑤人工筛选到的核酶能催化肽键的形成。⑥碎片反应。

证明大肠杆菌的转肽酶属于核酶的最有力证据来自 Harry F. Noller 设计的"碎片"反应（fragment reaction），Noller 利用它来测定转肽酶活性，即：

$$\text{CAACCA-fMet} + \text{嘌呤霉素} \longrightarrow \text{CAACCA-OH} + \text{fMet-嘌呤霉素}$$

其中 CAACCA-fMet 为 fMet-$tRNA_f^{Met}$ 在 3′端的片段，fMet 被 S^{35} 标记。在合适的条件下，碎片反应可由核糖体大亚基单独催化。Noller 发现，当使用蛋白酶 K 处理和 SDS 抽提后，大肠杆菌大亚基仍然保留 20%~40% 的转肽酶活性，而从一种嗜热菌得到的核糖体大亚基在使用相同的条件处理后能保留 80% 的转肽酶活性。然而，如果使用核糖核酸酶 T1 处理（它只水解 G 后面的磷酸二酯键，因此不会水解 CAACCA），大亚基的转肽酶活性就会完全丧失。以上实验的结果表明，23S rRNA 与转肽酶的活性直接有关，是一种名副其实的核酶。

转肽反应发生以后，新生的肽链连接到位于 A 部位的 tRNA 分子上，形成肽酰 -tRNA，P 部位则留下空载的 tRNA。在肽键形成的同时，A 部位上的 tRNA 受体茎发生旋转，以使新生肽链通过 P 部位进入大亚基上的离开通道。

3. 移位

经历了转肽反应以后，P 部位的 tRNA 已经成为空载的 tRNA（最初是空载的起始 tRNA，以后是空载的非起始 tRNA），空载的 tRNA 随后进入 E 部位，与此同时，移位（translocation）反应发生了。在 A 部位上形成的肽酰 -tRNA 连同与其结合的 mRNA 一起移动到 P 部位，从而为肽链延伸的下一轮循环做好了准备。注意在移位反应中，肽酰 -tRNA 上的反密码子与密码子的相互作用已不再是决定氨基酸特异性的因素，但是它对于维持移位反应的准确性（只移位一个密码子）以保持正确的可读框是至关重要的。

移位反应需要 EF-G 的帮助，EF-G 也是一种小 G 蛋白，其大小、形状以及电荷分布与结合有 EF-Tu 的氨酰 -tRNA 相似。EF-G·GTP 在 A 部位的附近与核糖体结合，促进移位反应的进行。EF-G·GTP 的结合可能将带有新生肽链的 tRNA 从 A 部位"推"到 P 部位，与此同时，空载的 tRNA 则从 P 部位转移到 E 部位。既然 tRNA 与 mRNA 通过密码子 / 反密码子的碱基配对结合在一起，mRNA 也就随之发生移位。一般认为移位反应是自发的，因为空载 tRNA 对 E 部位的亲和力要比对 P 部位高，而肽酰 -tRNA 对 P 部位的亲和力要比对 A 部位的亲和力高。但在移位过程中，GTP 并不发生水解，只是在移位完成以后，核糖体再次充当 EF-G 的 GAP，导致与 EF-G 结合的 GTP 水解成为 GDP 和 Pi。一旦 GTP 被水解，EF-G 立刻与核糖体解离，并依靠自身的一个结构域，作为 GEF 以再生出 EF-G·GTP。

每个细胞约含有 2×10^4 个 EF-G 分子，这与核糖体的数目相同。EF-G 和氨酰 -tRNA·EF-Tu 复合物在结构上惊人的相似，与核糖体的结合位点部分重叠，导致它们无法与核糖体同时结合（图 39-13 和图 39-14）。正是 EF-G 阻止氨酰 -tRNA 以及释放因子与 A 部位的结合，才保证了只有在移位反应结束后才可以进入下一轮循环。因此，只有在 EF-G 离开了核糖体，氨酰 -tRNA·EF-Tu 才能进入 A 部位，下一轮的进位反应才能发生。

在肽链延伸过程中，一个正在延伸的多肽链必须离开核糖体，结构研究表明，有一个狭窄的离开通道正为此用，多肽链在如此狭窄的空间内是很难发生折叠的。

Quiz6 细胞每合成一个肽键，至少需要消耗多少 ATP？

（四）多肽链合成的终止与释放

如图 39-15 所示，随着肽链的不断延伸，位于 mRNA 上的终止密码子最终进入 A 部位，由于没有相应的氨酰 -tRNA 的结合，RF1 或 RF2 便"有机可乘"。RF3 也是一种小 G 蛋白，它与 GTP 形成的复合物 RF3·GTP 可促进 RF1 和 RF2 的作用。一旦释放因子结合到 A 部位，核糖体上的转肽酶

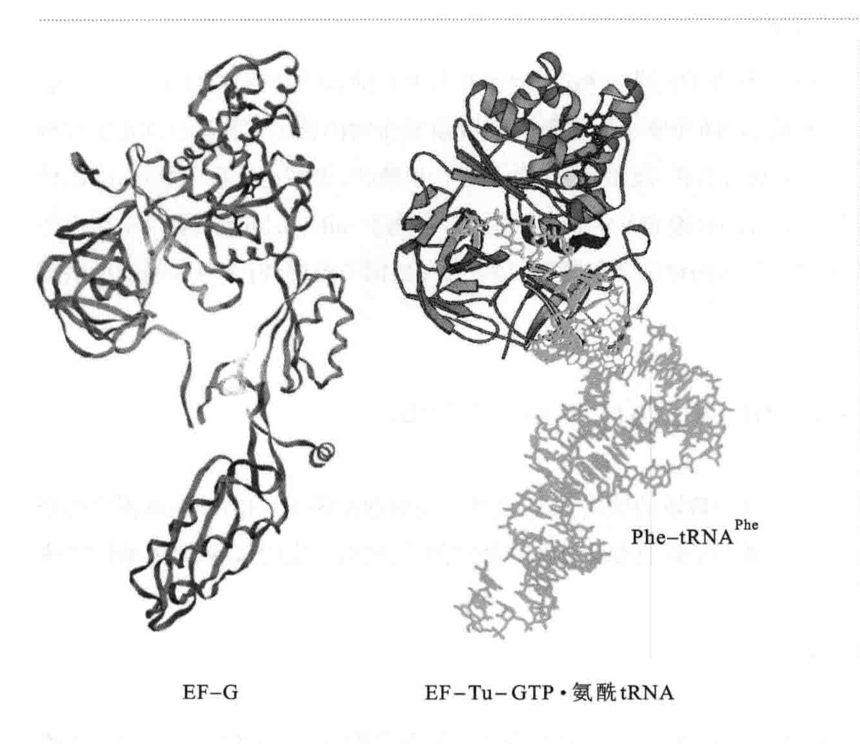


图 39-13　EF-G 与苯丙氨酰 -tRNA·EF-Tu 复合物在三维结构上的比较

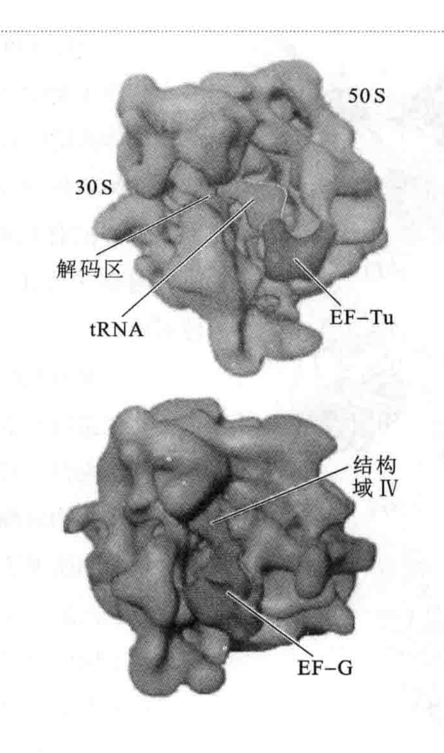


图 39-14　EF-Tu 和 EF-G 各自与核糖体结合的相互排斥

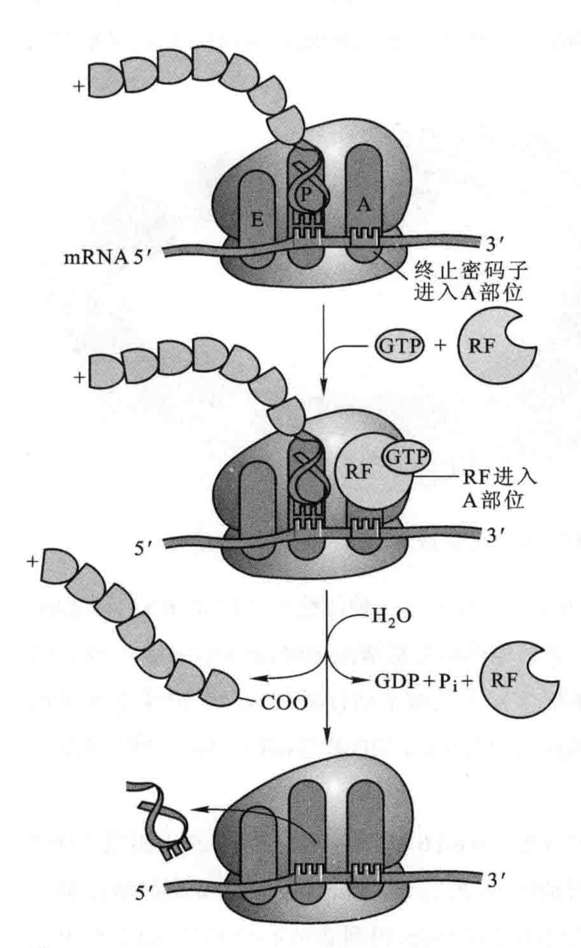


图 39-15　细菌多肽链合成的终止与释放

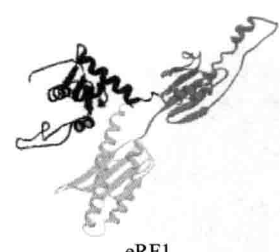
在活性上就会发生改变，转而催化肽酰基转移给水分子，导致肽酰 -tRNA 的水解，肽链因此被释放出来。随后空载的 tRNA 离开核糖体，释放因子因为 RF3 的 GTP 酶活性将结合的 GTP 水解而得以释放。最后，在核糖体循环因子（RRF）的帮助下，核糖体与 mRNA 解离，解离的核糖体进入下一轮翻译。

核糖体最重要的功能也许不是催化肽键的形成，而是如何防止肽酰 -tRNA 的酯键提前发生水解，因为肽键形成在热力学上是一个很容易发生的反应。显然，将水分子排除在转肽酶活性中心之外是防止肽链提前释放的关键。释放因子很可能使用分子模拟（molecular mimicry）的方法，将自己“乔装打扮”成氨酰 -tRNA 的结构（图 39-16），并将与它结合的水分子带入转肽酶的活性中心，随后在核糖体的协助下，促进水分子进攻酯键，导致肽链的释放。在多肽链释放以后，mRNA 和空载 tRNA 还短暂地结合在 70S 核糖体上（tRNA 留在 P 部位），直到在 RRF 和 EF-G·GTP 的作用下才发生解离。

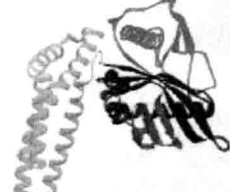
RRF 是细菌翻译必需的蛋白质因子，它和 EF-G 一起参与最后翻译系统的解体。一旦 GTP 水解，RRF 和 EF-G 即可导致核糖体解离成亚基。随后，在 IF3 的帮助下，空载的 tRNA 离开 30S 亚基。RRF 的作用方式与 EF-Tu 和 EF-G 一样，通过模拟 tRNA 的形状与核糖体上的 tRNA 结合位点结合。

二、真核生物的细胞质翻译系统

真核生物的翻译过程与细菌在许多方面非常相似，但也有不少差别，具体反映在以下9个方面：①核糖体的结构有所不同；②细菌的翻译和转录是紧密偶联，而真核生物的转录和翻译分别发生在细胞核和细胞质，因此两者不存在偶联关系；③起始tRNA并不进行甲酰化，也不能进行甲酰化；④起始密码子的识别采取完全不同的机制；⑤起始tRNA与小亚基的结合先于mRNA；⑥起始阶段不仅需要GTP，还需要ATP；⑦起始因子种类与结构比细菌要复杂得多；⑧释放因子有两种；⑨对抑制剂的敏感性不同。



eRF1



RF2



RRF

图 39-16 eRF1、RF2 和 RRF 三维结构的比较

e39-2 真核生物参与翻译的起始因子、延伸因子和终止因子

（一）氨基酸的活化

与细菌没有多少差别，所不同的是 Met-tRNA$_i^{Met}$ 并不进行甲酰化。

（二）翻译的起始

真核生物的翻译的起始可分为4个阶段的反应，至少需要十余种起始因子（eIF）或辅助蛋白的帮助。这些起始因子和参与延伸的延伸因子以及参与终止的释放因子，按照一定的次序作用，共同完成一条多肽链的合成。

各个阶段的反应分述如下：

1. mRNA的准备和检查

由于真核系统的mRNA要经历复杂的后加工，因此翻译系统首先需要对mRNA进行检查，以确保只有加工完好的mRNA才能用作模板。参与这一步反应的起始因子为eIF4系列，其中eIF4E为帽子结合蛋白，专门与mRNA的5′端的帽子结合，eIF4G是一种接头分子，既能与eIF4E结合，又能与结合在3′端尾巴上的PABP结合，还能结合eIF3，使mRNA的5′和3′端在空间上相互靠近成环（图39-17）。mRNA的环化能很好地解释多聚A尾巴为什么能提高翻译的效率：一旦核糖体完成翻译通过多聚A成环的mRNA，重新释放的核糖体亚基所处的位置恰到好处，非常适合在同一个mRNA分子上重新启动翻译。

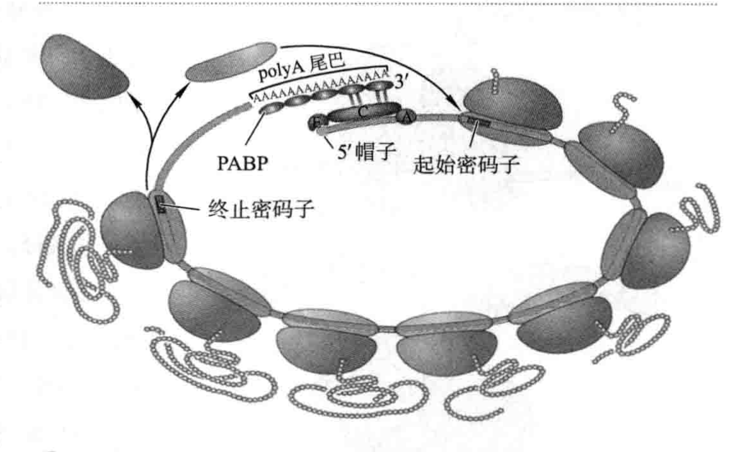


图 39-17 真核生物 mRNA 成环模型（Watson 等，2007）

eIF-4G与PABP的结合不仅保证了只有成熟的完整的mRNA才被翻译，还将其他起始因子招募进来，例如eIF4A和eIF4B。eIF4A是一种依赖于ATP的RNA解链酶，负责破坏mRNA 5′端的二级结构，暴露起始密码子，为核糖体的扫描清除障碍，而eIF4B是一种RNA结合蛋白，其功能是刺激eIF4A的解链酶活性。eIF4F实际上是帽子结合蛋白eIF4E、依赖于ATP的解链酶eIF4A和eIF4G的复合物，这三种蛋白之间的相互作用加强了eIF4E与mRNA帽子结构的结合。

2. 43S预起始复合物的形成

这一步反应涉及的起始因子有eIF1、eIF1A、eIF2、eIF3和eIF6（图39-18）。其中，eIF1相当于细菌的IF1。eIF2是一种小G蛋白，只有与GTP结合才有活性，它相当于细菌的IF2，负责识别和结合Met-tRNA$_i^{Met}$。eIF3与40S小亚基结合，阻止60S大亚基与小亚基的结合，但刺激Met-tRNA$_i^{Met}$·eIF2·GTP与小亚基的结合。

3. 起始密码子的识别及48S起始复合物的形成

43S预起始复合物和eIF5附着到已检查合格的mRNA分子的5′端（图39-19），在eIF4A、eIF4B和eIF4F的帮助下，mRNA的5′-UTR发生解链，预起始复合物从5′端向3′端进行扫描，直到发现第一个AUG，通常就以这个AUG作为起始密码子。

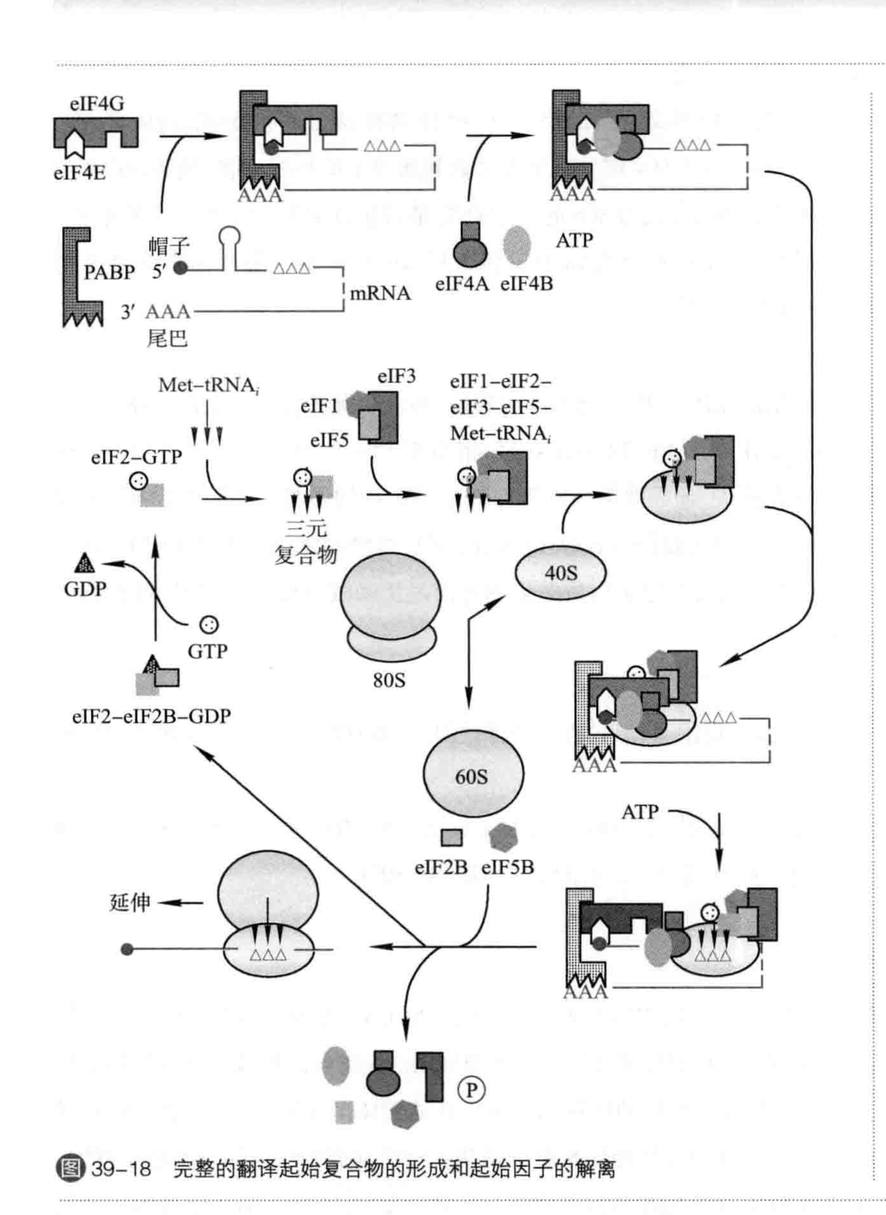


图39–18 完整的翻译起始复合物的形成和起始因子的解离

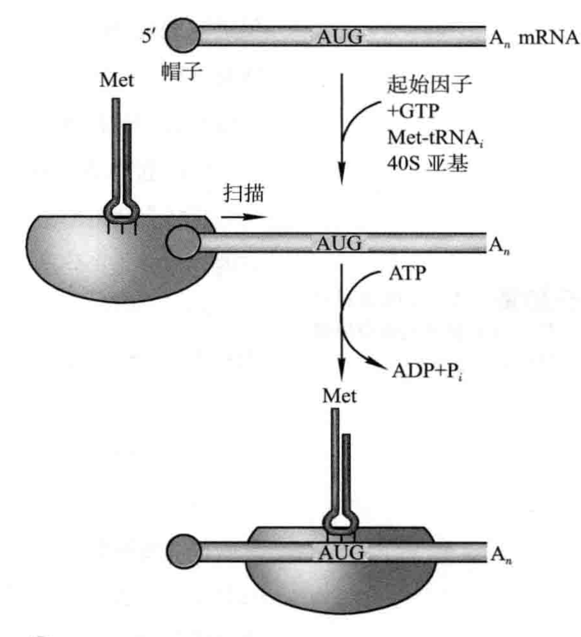


图39–19 真核翻译系统发现起始密码子的扫描机制

然而，5%~10% 的 mRNA 并不使用第一个 AUG 作为起始密码子，这是因为第一个 AUG 所处的序列环境对能否选做起始密码子有很大影响。Marilyn Kozak 所做的研究表明，在一致序列 CCRCC**AUG**G（R 代表嘌呤碱基）之中的 AUG 起始的效率最高。如果嘧啶碱基占据 –3 或 +4 区，起始复合物 tRNA·40S 就会漏过此 AUG，继续扫描，直到发现新的 AUG。小亚基在进行扫描的时候，需要 eIF4A 的解链酶活性，以解除 mRNA 5′端的二级结构。

Quiz7 你如何解释真核生物的蛋白质基因必须以单顺反子的形式存在，而原核生物的蛋白质基因既可以以单顺反子，也可以以多顺反子的形式存在？

在扫描中，起始 tRNA 通过反密码子 / 密码子的碱基配对，来进行校对以发现 AUG，eIF2 的 GTP 酶活性则充当计时器，允许有足够的时间校对。

真核翻译系统还有一种识别起始密码子的机制叫内部进入（internal entry），使用这种机制的主要是细小核酸病毒（picornavirus），如脊髓灰质炎病毒。这一类病毒的 mRNA 在 5′端无帽子结构，但有很长的 5′UTR（600~1 000 nt）。5′UTR 内有多个 AUG，但都不能作为起始密码子。通过对 5′UTR 进行的缺失突变分析发现，在 5′端起始密码子的周围存在一种内部核糖体进入位点（internal ribosome entry site，IRES），形成特殊的二级结构和三极结构，核糖体可以识别它，并从这个区域直接进入然后发现周围的起始密码子。

Quiz8 如何能够让一种共价闭环的 mRNA 也能被翻译出蛋白质？

显然，使用内部进入的翻译起始除不需要帽子结合蛋白 eIF4E 以外，需要其他所有的起始因子，其中的 eIF4G 可以直接作用于 IRES，也可能通过细胞内的 IRES 结合蛋白间接作用。结合在 IRES 的 eIF4G 再将 eIF3 和 40S 亚基招募上去。

4. 起始因子的解离与 60S 大亚基的结合

在起始密码子被识别以及 48S 复合物形成以后，eIF5、eIF5B 便刺激 eIF2 的 GTP 酶活性，促进与 eIF2 结合的 GTP 水解。与 eIF2 结合的 GTP 的水解导致绝大多数起始因子的释放，同时激活 eIF5B 的 GTP 酶活性。于是，GTP 受到 eIF5B 的水解，成为 80S 起始复合物是否正确装配的最后一个检查点。当与 eIF5B 结合的 GTP 被水解以后，剩余的所有起始因子都离开 40S 小亚基。随后，60S 大亚基与 40S 小亚基结合，形成完整的翻译起始复合物。

（三）翻译的延伸

真核生物的肽链延伸反应与细菌相似，同样不断地经历进位、转肽和移位反应，只是由 eEF-1 代替了 EF-Tu 和 EF-Ts，其中 eEF-1α 相当于 EF-Tu，eEF-1βγ 相当于 EF-Ts，eEF-2 代替了 EF-G，转肽酶的活性由 28S rRNA 提供。在真菌中，还需要第三种延伸因子 eEF-3 的参与。eEF-3 的功能是能促进空载的 tRNA 从 E 部位释放，从而刺激氨酰 -tRNA 进入 A 部位，这种功能依赖于它的 ATP 酶活性对 ATP 的水解。eEF-1 和 eEF-2 可被特定的蛋白质激酶催化而经历磷酸化修饰，这种修饰在体内可能具有调节延伸反应的功能。

（四）蛋白质合成的终止

真核生物的肽链终止反应与细菌也很相似，但只有 2 个释放因子，而且没有 RRF，由 eRF1 识别所有的 3 种终止密码子。

Quiz9 原核生物与真核生物体内参与翻译的蛋白质哪些属于 G 蛋白？

eRF1 的作用也使用分子模拟的手段，其三维结构与 tRNA 相似。但 eRF1 并不单独作用，需要和 eRF3 形成异源二聚体。eRF3 也是一种 G 蛋白，其功能相当于细菌的 RF3。

三、细胞器翻译系统

线粒体和叶绿体是两种半自主细胞器，在它们的基质内不仅发生 DNA 复制和转录，还进行翻译。这两种细胞器翻译的特征似乎更接近细菌，如核糖体的结构更像细菌。大肠杆菌和线粒体的翻译系统的交换实验表明，大肠杆菌的 tRNA 和参与翻译的各种辅助蛋白在线粒体翻译系统是有功能的，这说明这两种翻译系统在功能上是等同的。此外，两种细胞器翻译系统对抑制剂的敏感性也类似于细菌。参与起始、延伸和终止反应的各蛋白质因子与细菌对应的成分具有高度的同源性，第一个被参入的氨基酸通常是甲酰甲硫氨酸（哺乳动物的线粒体除外）。但是，线粒体翻译系统和叶绿体系统也各有自己特有的性质。以线粒体翻译系统为例：首先，与大肠杆菌的核糖体相比，线粒体核糖体内的 rRNA 含量减半，有的丧失了 5S rRNA，但蛋白质含量有所增加。在 rRNA 上有结合位点的核糖体蛋白质缩短了或丢失了。其次，它的 tRNA 极为简练。与细菌和真核细胞质翻译系统内的 50 种以上的 tRNA 相比，动物线粒体仅有 22 种 tRNA。22 种 tRNA 要识别 60 种密码子，必须遵守更为宽松的摆动法则。线虫的线粒体翻译系统有 20 种 tRNA 无 T 臂，余下两种 $tRNA^{Ser}$ 缺乏 D 臂。有趣的是，线虫的线粒体有 2 种 EF-Tu（EF-Tu1 和 EF-Tu2），EF-Tu1 识别缺乏 T 臂的 tRNA，而 EF-Tu2 识别缺乏 D 臂的 $tRNA^{Ser}$，这 2 种 EF-Tu 尚没有在其他生物体内发现。EF-Tu1 在其 C 端多出一段由 57 个氨基酸残基组成的结构域，该结构域似乎是用来弥补缺失的 T 臂，以形成正常的氨酰 tRNA·EF-Tu1·GTP 三元复合物。此外酵母细胞线粒体内的翻译还依赖于由核基因编码的某些蛋白质因子的激活。事实上，线粒体翻译系统除了 rRNA 和 tRNA 是由线粒体 DNA 编码的以外，其他所有的参与翻译的蛋白质都是由核基因编码。最后，线粒体翻译系统使用的密码子与标准密码子有所不同。

就叶绿体翻译系统而言，它在很多方面与细菌很相似，但是，在叶绿体中只有三分之一的 mRNA 在 5′-UTR 含有 SD 序列，而且该序列对翻译的起始并不是绝对必需的。已有证据表明，叶绿体可能存在多种翻译起始机制，其中许多 mRNA 的翻译与线粒体相似，也需要核基因编码的蛋白质的激活。叶绿体翻译系统有一套完整的 rRNA，但也只有 30 种 tRNA。在某些叶绿体内，有的 mRNA 通过编辑（主要是 C→U）创造起始密码子、终止密码子或者改变密码子以维护保守的氨基酸残基的稳定。

四、古菌的翻译系统

正如DNA复制和转录一样，古菌的翻译与真核生物也十分相似。这表现在以下几个方面：①虽然核糖体的大小与细菌一样，但构成核糖体的rRNA和蛋白质与真核生物的亲缘关系更密切。真核生物的核糖体含有78种蛋白质和4种rRNA。其中有34种蛋白质同时存在于细菌和古菌，还有33种也存在于古菌，剩下的11种为真核生物特有的，没有任何一种只存在于真核生物和细菌。②参与翻译各个阶段的辅助蛋白质因子的数目以及结构的同源性接近真核生物，但各种因子的组成倾向于由单个亚基组成，而不像真核生物由多个亚基组成。例如参与翻译起始阶段的起始因子，有的既存在于细菌、真核生物和古菌，有的只存在于古菌和真核生物，但没有哪一种仅存在于细菌和古菌。再如，在翻译终止阶段，细菌需要RRF，但古菌和真核生物都不需要RRF。③第一个参入的氨基酸和真核生物一样，都是甲硫氨酸，而不是细菌的甲酰甲硫氨酸。④对抑制剂，特别是对抗生素的敏感性相似。例如，真核生物的细胞质翻译系统和古菌翻译系统的核糖体移位反应都受到白喉毒素的抑制，但都不受氯霉素、链霉素和卡拉霉素的抑制。

然而，古菌的翻译系统与细菌也有某些特征很相似。例如：没有5.8S rRNA，mRNA无帽子结构，多为多顺反子，有SD序列，翻译与转录是偶联的。

第四节　mRNA的质量控制

细胞内的mRNA并不总是正常的，某些因素会导致一个mRNA分子在其ORF内提前出现终止密码子（premature stop codon，PTC）或者无终止密码子（nonstop），由这些异常的mRNA翻译出来的蛋白质可能是没有功能的，甚至是有害的。为了及时清除异常mRNA对细胞可能产生的危害，无论是原核生物，还是真核生物，都具有特定的质量控制系统，来实时监控细胞内的mRNA，以及时处理各种异常的mRNA。但两类生物控制的方法是不同的，像细菌主要是对异常的mRNA进行抢救翻译（rescuing translation），而真核生物直接对异常的mRNA进行降解。

细菌mRNA一般很快水解，其半衰期较短，因此，mRNA有很大的可能性丢掉了它的3′端。倘若这种情况真的发生了，后果就很可能非常严重。不妨设想，假定一个mRNA分子因此或者因突变丢失了它的终止密码子，那么将没有终止信号促进核糖体的解离。任何与这种有缺陷的mRNA结合的核糖体当翻译到断裂的末端，将裹足不前，难以解离。

面对上述情况，细菌已发展了一套专门的机制，来处理这些在翻译上有缺陷的mRNA，这种机制就是抢救翻译，而抢救翻译的手段是反式翻译（*trans*-translation）。反式翻译不同于通常形式的顺式翻译，它由两个mRNA翻译成一条融合的肽链，其中一个mRNA分子无终止密码子，另一个mRNA分子是tmRNA。

tmRNA由349~411 nt组成，广泛存在于各种细菌，也存在于某些生物的叶绿体和线粒体内，但并不存在于古菌和真核生物的核基因组。绝大多数tmRNA是一个连续的RNA分子，少数tmRNA由2个独立的RNA分子通过碱基配对组装而成。大肠杆菌的tmRNA也被称为10SaRNA，含有363 nt，由457 nt的前体经后加工而成。

tmRNA可分成2个部分（图39-20）：第一部分包括5′端（约50个核苷酸）和3′端（约70个核苷酸）的核苷酸，两者之间形成一段长长的配对区域，3′端的最后三个核苷酸序列为CCA。在不同的tmRNA分子之中，第一部分的结构十分相似，都折叠成类似tRNA的结构，因此它的功能是作为一种tRNA来用。所有tmRNA的受体茎上的第三个碱基对都是GU碱基对，这正是细菌$tRNA^{Ala}$的个性，因此它可以携带Ala；第二部分由内部的核苷酸组成，包括2个茎环结构和4个假节结构，这一部分差别很大，但都有一个潜在的小ORF，因此它的功能是作为一种mRNA，编码一段寡肽序列。

如图39-21所示，面对缺乏终止密码子的mRNA，核糖体照样能够结合上去，起始多肽链的合成，并持续到mRNA的3′端。但由于无终止密码子，肽酰-tRNA并不能释放，而是一直与P部位结合，直

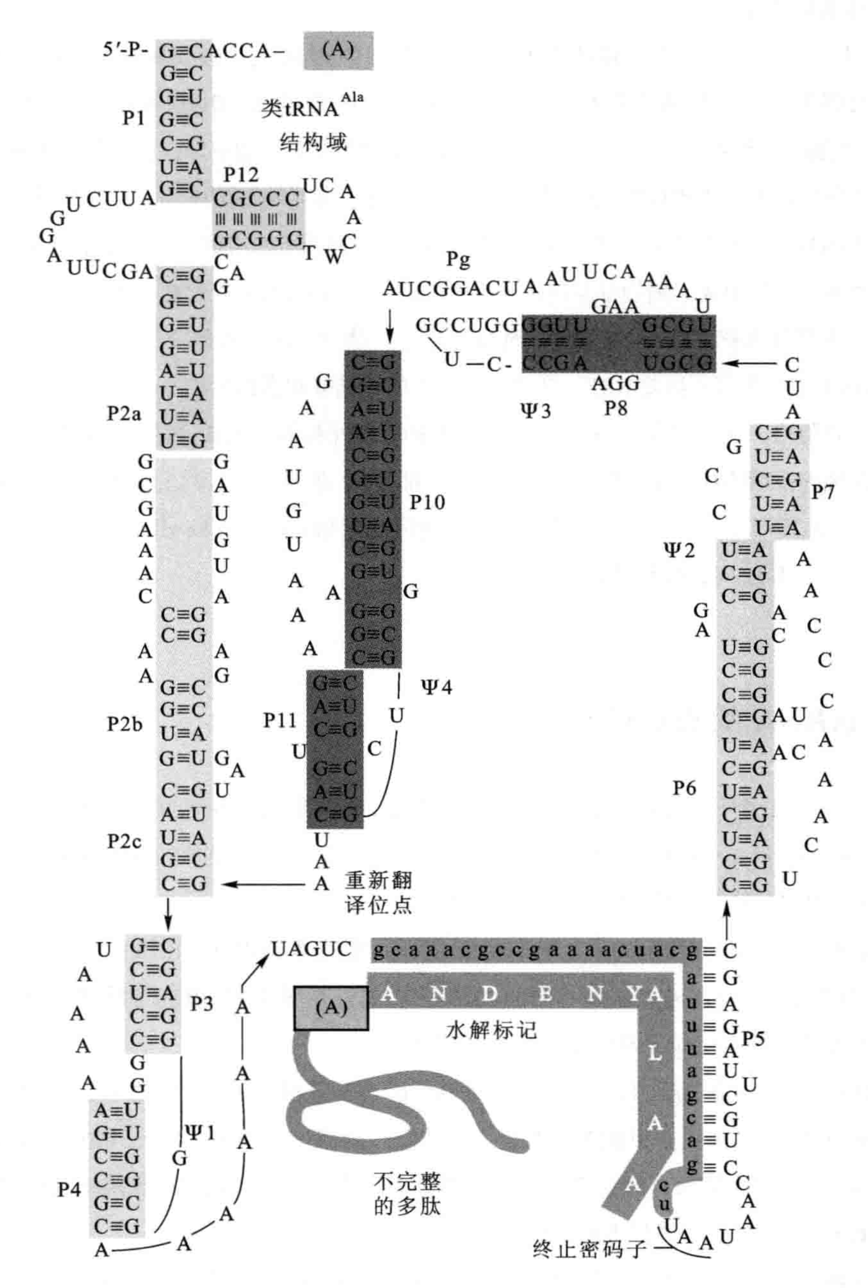


图 39-20 大肠杆菌 tmRNA 的结构

到在 SmpB 蛋白和 EF-Tu·GTP 的帮助下，Ala-tmRNA 进入核糖体的 A 部位。与氨酰 -tRNA 进入 A 部位有所不同的是，它不可能通过密码子与反密码子的相互作用。有证据表明，tmRNA 有一段序列与 16S rRNA 上的一段序列互补，tmRNA 很可能通过这种互补关系才结合到 A 部位上去的。在转肽酶的催化下，肽酰基被转移到 Ala-tmRNA 上的 Ala 的游离氨基上，随后核糖体照常移位，并以 tmRNA 上的 ORF 的后 10 个密码子序列接着充当模板，在翻译出 10 个氨基酸残基（ANDENYALAA）以后，就遇到了终止密码子，肽链合成得以正常终止，而释放出来的肽链是一个融合蛋白，在 C 端含有一段 11 肽序列。这段 11 肽序列的最后两个氨基酸残基是一种降解信号，可被胞内的一些特殊的蛋白酶（ClpAP 和 ClpXP）识别，最终导致这些被抢救翻译出来的异常多肽能及时水解。

由此可见，使用 tmRNA 既解除了细胞在翻译无终止密码子 mRNA 时核糖体不能解离的问题，又能保证翻译出来的异常多肽能被及时清除，防止它们在细胞内堆积，产生毒害。

真核生物有 2 种高度保守的机制，专门处理异常的 mRNA：一种是无义介导的 mRNA 降解（nonsense-mediated mRNA decay，NMD）——专门处理含有 PTC 的 mRNA，另一种则是无终止 mRNA 降解（nonstop mRNA decay，NSD）——专门处理无终止密码子的 mRNA。

e39-3 NMD 和 NSD

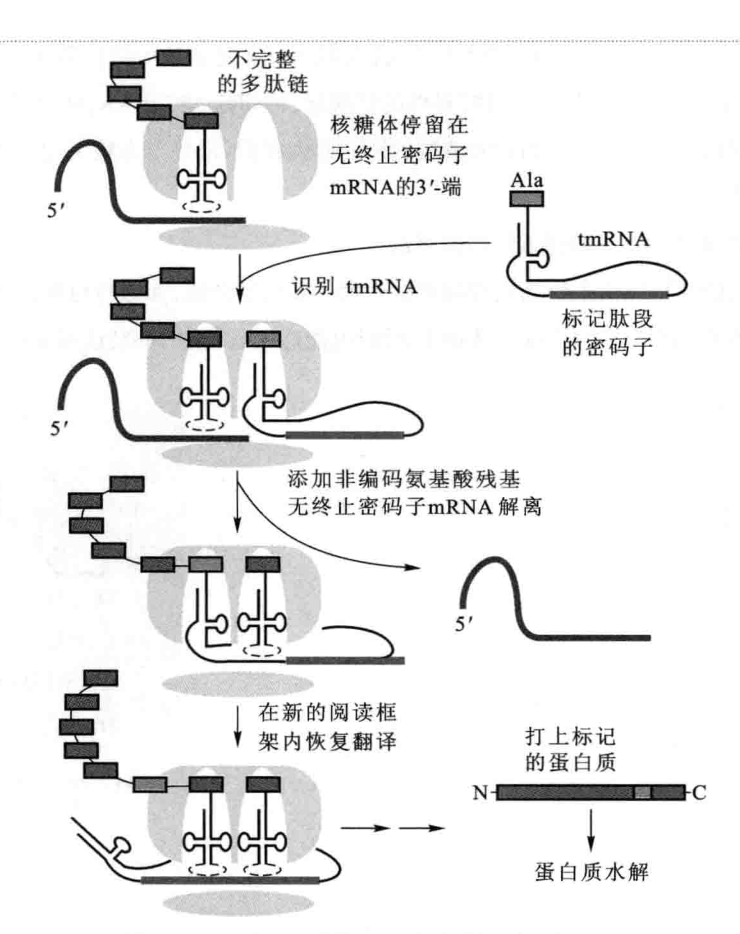


图39-21 使用tmRNA的反式翻译

第五节 翻译的抑制剂

翻译是维持细胞的正常功能所必需的，许多抑制剂能够抑制参与翻译的某一重要成分的活性，从而影响到被作用的细胞的生存。自然界有各种各样的翻译抑制剂，绝大多数作用位点是核糖体，也有的作用位点是起始因子或延伸因子等。根据抑制的对象，抑制剂通常可以分为3类：

1. 细菌翻译系统的抑制剂

这一类抑制剂大多数是人们熟悉的抗生素类药物。例如链霉素（streptomycin）、氯霉素、林可霉素（lincomycin）、稀疏霉素（sparsomycin）、黄色霉素（kirromycin）、红霉素（erythromycin）和四环素（tetracycline）等，作用对象是细菌翻译系统，真核生物则不敏感。但是，如果用量过大，细胞器翻译系统也会受到抑制，在这种情况下，真核细胞的功能同样受到影响，这是许多抗生素产生副作用的主要原因。

链霉素是一种氨基糖苷类（aminoglycoside）抗生素，在低浓度下，能导致核糖体误读mRNA，这时它只会抑制敏感菌的生长，但不会杀死敏感菌。但在高浓度下，链霉素则完全抑制了翻译起始，敏感菌会被杀死。氯霉素是一种广谱抗生素，它与林可霉素和稀疏霉素一样与核糖体50S亚基结合，抑制大亚基的转肽酶活性。四环素也是一种广谱抗生素，它与小亚基结合而抑制氨酰-tRNA结合，它也能抑制ppGpp的合成而解除严紧反应。红霉素作用位点是50S亚基上的多肽离开通道，阻断正在生长的肽链的离开，从而阻滞翻译。

2. 真核翻译系统的抑制剂

这类抑制剂只抑制真核翻译系统，不抑制细菌翻译系统，例如白喉毒素（diphtheria toxin）、蓖麻毒素（ricin）、放线菌酮（cycloheximide）、茴香霉素（anisomycin）和α-帚曲霉素（α-sarcin）。其中，白喉毒素催化eEF-2的ADP-核糖基化修饰，从而导致移位反应不能发生，而蓖麻毒素遇到80S核糖体，可切

下 28S rRNA 上的一个腺嘌呤，导致核糖体失活，放线菌酮只抑制真核生物核糖体大亚基的转肽酶活性。α- 帚曲霉素在真核细胞内作为一种特异性的核酸酶，可切断 28S rRNA，从而导致氨酰 -tRNA 与核糖体的结合受到抑制。茴香霉素和放线菌酮的抑制原理相似，都是与真核生物核糖体大亚基结合，抑制其转肽酶活性。

Quiz10 与其他抗生素相比，为什么要达到相同的抑制效果嘌呤霉素需要更高的剂量？

3. 既抑制原核又抑制真核翻译系统的抑制剂

少数抑制剂既能抑制原核生物又能抑制真核生物的蛋白质合成，如嘌呤霉素，它的分子结构与酪氨酰 -tRNA 非常相似（图 39-22），因此在翻译的时候也能进入 A 部位，而转肽酰酶照样把 P 部位上的

嘌呤霉素　　酪氨酰 - tRNA

图 39-22　嘌呤霉素的化学结构

框 39-2　身边的生物化学——为什么蓖麻毒素不会杀死蓖麻自身？

2003 年，伦敦警察突袭了一处公寓，逮捕了一些恐怖分子嫌疑人，发现了实验用的玻璃器皿和蓖麻子。这次行动提醒了世人，恐怖分子可能一直在想利用蓖麻毒素作为生化武器。蓖麻毒素的毒性是很强的，仅次于钚和肉毒杆菌毒素，其毒性是氰化物的 6 000 倍，是眼镜蛇毒液毒性的两倍，1 g 蓖麻毒素可致 3 600 人死亡。蓖麻毒素存在于蓖麻植物所有的组织，但在种子中的含量最高。

蓖麻毒素的毒性强是因为它在进入真核细胞之后作为一种核糖体失活蛋白，使核糖体失活，从而导致蛋白质合成受到强烈抑制。蓖麻毒素是由 2 条肽链组成，一条肽链被称为 RTB，具有凝集素的活性可与细胞膜结合，另一条肽链被称为 RTA，在 RTB 与膜结合以后进入细胞内，作为一种高度特异性的 28S rRNA 糖苷酶，去水解 28S rRNA 的 A4324，由于 A4324 处于 28S rRNA 一段高度保守的十二聚核苷酸序列之中，这个位置恰好是 eEF-2 与核糖体结合的地方，因此，一旦 A4321 被水解下来，核糖体就不能催化蛋白质合成的移位反应，从而导致翻译受阻。蓖麻毒素作用的效率是很高的，1 分子蓖麻毒素在 1 min 内可以让 15 000 个核糖体失活。由于蓖麻毒素的作用位点的保守序列几乎存在于所有真核生物核糖体的 28S rRNA 分子上，因此，真核生物的蛋白质合成都会受到它的抑制，也包括蓖麻植物本身。但为什么蓖麻毒素并没有杀死蓖麻自身呢？

原来是在种子形成的时候，蓖麻毒素就像胰岛素一样，是由单一的 mRNA 翻译而成的，这个单一的 mRNA 编码的是蓖麻毒素的前体——前蓖麻毒素原（preproricin）。前蓖麻毒素有 576 个氨基酸残基，其中 N 端前 35 个残基是它进入内质网腔的信号序列，在前体进入内质网腔以后被切除，产生蓖麻毒素原。蓖麻毒素原仍然由一条肽链组成，因为 RTA 和 RTB 之间由 12 个残基组成的连接肽还没有除掉。蓖麻毒素原会在内质网腔发生 N- 糖基化修饰，同时进行折叠并形成 5 个二硫键。其中有 4 个二硫键在 RTB 内部，1 个是在 RTA 的 C 端和 RTB 的 N 端之间。折叠好的蓖麻毒素原进入高尔基体，再通过小泡运输，进入液泡，在液泡中被一种半胱氨酸蛋白酶去除 RTB 和 RTA 之间的 C 肽，而变成成熟的形式并被分泌到细胞壁。由此可见，蓖麻毒素不会让蓖麻自身中毒，有两个原因：一是最初制造出来的是没有活性的前体，二是等它最后有活性的时候，已经进入了液泡，被隔离起来，无法回到细胞质去作用其中的核糖体。

肽酰-tRNA 的肽酰基转移到它的氨基上，形成肽酰嘌呤霉素，然而形成的肽酰嘌呤霉素并不能移位，而是很快与核糖体解离，而一旦解离，肽链合成便提前结束。

需要注意的是，古菌对绝大多数抑制剂的敏感性与真核生物相似，如白喉毒素。

第六节 蛋白质在细胞内的降解

细胞内的蛋白质在不断地合成，同时也在不断地降解。这样的过程看似浪费，其实可以为细胞清除一些不正常的蛋白或者积累过多的酶和调节蛋白，从而使细胞代谢井然有序。然而，不同性质的蛋白质的半衰期差别很大，从几分钟到几百个小时不等。通常承担重要调节作用的蛋白质，如代谢反应中的限速酶、癌基因的产物和细胞周期蛋白的半衰期最短，而由管家基因编码的管家蛋白的半衰期最长。

与蛋白质在消化道内进行的不受控制的、也完全不需要 ATP 的降解途径相比，细胞内发生的蛋白质降解是受到严格调控的，其中有的降解过程需要 ATP，有的不依赖于 ATP。

一、细胞内不依赖于 ATP 的蛋白质降解系统

在真核细胞中，主要存在 2 套蛋白质降解系统：一套是溶酶体系统，另一套是蛋白酶体(proteasome)系统，其中溶酶体降解系统不依赖于 ATP。

溶酶体降解系统主要负责处理由胞外摄入到胞内的蛋白、在受体介导的内吞中进入内吞小泡的细胞膜蛋白和被自体吞噬体（autophagosome）吞噬的胞内蛋白质。

二、细胞内依赖于 ATP 的蛋白质降解系统

真核细胞内依赖于 ATP 的蛋白质降解主要是受泛素介导的蛋白酶体降解系统。该系统不仅需要消耗 ATP，还需要一种叫泛素（ubiquitin，Ub）的蛋白质，而降解的真正场所是在蛋白酶体内。蛋白酶体是一种完全由蛋白质组成的细胞器，存在于细胞质和细胞核，负责处理内源的蛋白质，主要包括转录因子、周期蛋白、病毒编码的蛋白质、基因突变或翻译异常产生的错误折叠的蛋白质以及细胞质中受到损伤的蛋白质。故这种降解系统对于基因表达和细胞周期的调控，以及细胞面对各种胁迫因素的反应等方面都具有非常重要的意义。

（一）泛素的结构与功能

泛素广泛存在于古菌和所有的真核生物，但不存在于细菌。它由 76 个氨基酸残基组成，是一种高度保守的蛋白质，酵母与人的泛素在一级结构上只差 3 个氨基酸残基。在三维结构上，泛素是一个结构紧密的球蛋白，但其 C 端四肽序列（LRGG）却离开蛋白主体伸向水相，这有助于它与其他蛋白质形成异肽键(图 39-23)。

泛素的作用需要消耗 ATP，以特异性识别需要降解的靶蛋白。泛素本身并不降解蛋白质，它仅仅是给降解的靶蛋白加上标签，降解过程由 26S 蛋白酶体执行。

（二）泛酰化反应及其功能

泛素的功能是参与蛋白质的泛酰化反应。泛酰化由 3 步反应组成(图 39-24)，依次由泛素活化酶（ubiquitin-activating enzyme，E1）、泛素结合酶（ubiquitin-conjugating enzyme，E2）和泛素连接酶（ubiquitin ligase，E3）催化。首先，Ub 以依赖于 ATP 的方式被 E1 激活（反应机理类似于脂肪酸和氨基酸的活化），然后，E2 和 E3 一起识别靶蛋白并催化泛素 C 端的 Gly76 的羧基与靶蛋白分子上的赖氨酸残基的 ε-NH_2 形成异肽键，导致靶蛋白的泛酰化。

细胞有多种类型的 E2，不同的 E2 行使不同的功能。例如，同属于 E2 的 Ubc2 和 Ubc3 分别参与 DNA 修复和降解周期蛋白。同样，也发现多种类型的 E3，它们具有不同底物特异性，其中有一类被用

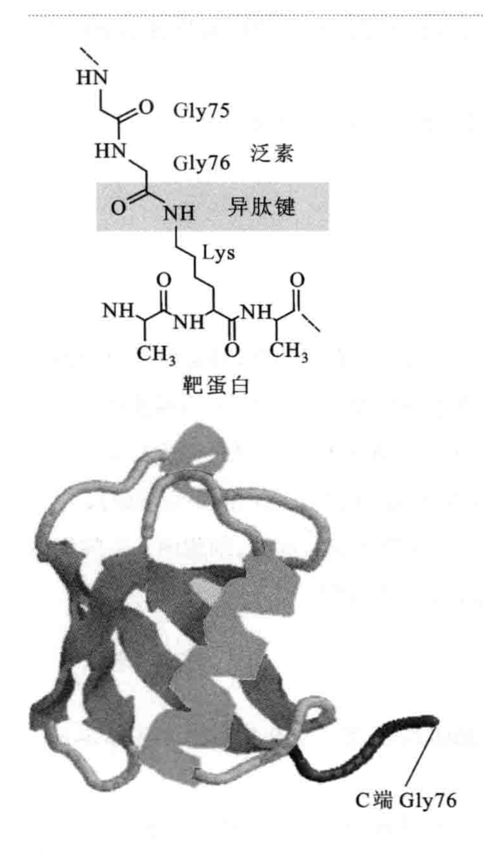


图 39-23 泛素的三维结构及其与靶蛋白形成的异肽键

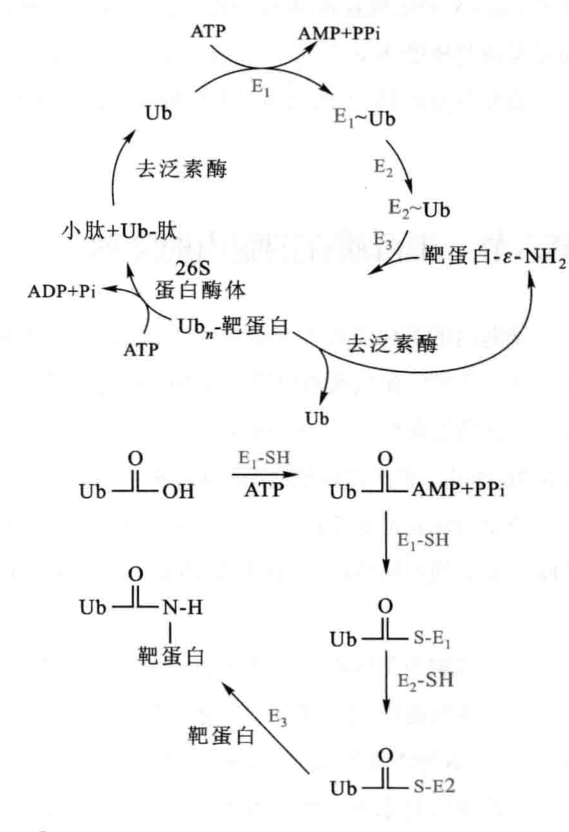


图 39-24 蛋白质泛酰化与蛋白质的定向水解

来识别 N 端规则(见下),还有的针对特定的底物蛋白。

蛋白质的泛酰化有单泛酰化(monoubiquitylation)、多重单泛酰化(multiple monoubiquitination)和多聚泛酰化(polyubiquitination)3 种形式。如果靶蛋白上仅连接 1 个泛素分子,就称为单泛酰化;如果同一个靶蛋白分子上有几个 Lys 残基连接上单个泛素分子,就称为多重单泛酰化;如果连接到靶蛋白上的泛素分子本身的 Lys 残基(Lys48、Lys63 或 Lys29)又发生泛酰化,以致多个泛素分子以异肽键连接在一起,就称为多聚泛酰化。由于泛素的 Lys48 残基的侧链从蛋白主体伸出到水相,因此,它特别适合作为 1 个泛素分子与另 1 个泛素 Gly76 形成异肽键的 ε- 氨基的供体。

Quiz11 总结一下在细胞里哪些分子可以和 Lys 残基的 ε- 氨基共价结合?

现在已经很明确,不同形式的泛酰化具有不同的功能,只有发生在泛素 Lys48 残基上的多聚泛酰化,才参与在蛋白酶体内的蛋白质降解,实际上,多泛素链的形成是将需要降解的蛋白质打上"死亡标签",它似乎是靶蛋白进入蛋白酶体进行降解的先决条件。而通过 Lys63 或 Lys29 的多聚泛酰化参与细胞其他功能,包括 DNA 修复和细胞内吞。

(三) 泛酰化反应的信号

蛋白质在降解之前,细胞应该具有某种识别机制,以区分需要水解和不需要水解的蛋白质。对于通过泛酰化介导的蛋白酶体降解途径降解的蛋白质来说,首先需要被打上泛素标记,那么是什么因素决定一个蛋白质是否该打上"该死"的泛素标记呢?已有证据表明,在蛋白质分子上,有一些特殊的信号可供泛素 - 蛋白酶体降解系统识别,这些信号有:

1. N 端规则

Alexander Varshavsky 在 1986 年已发现,决定一个蛋白质半衰期的重要因素是其 N 端的氨基酸的性质。某些氨基酸在 N 端能延长蛋白质的半衰期,而某些氨基酸在 N 端能缩短蛋白质的半衰期(表 39-7)。一个蛋白质半衰期与其 N 端氨基酸性质之间的这种关系被称为 N 端规则(N-end rule),这个规则适用于原核生物和真核生物。

表 39-7 N端氨基酸与蛋白质半衰期之间的关系

N端残基		半衰期
稳定性氨基酸残基	Met、Gly、Ala、Ser、Thr、Val	>20 h
去稳定的氨基酸残基	Ile、Gln	约 30 min
	Tyr、Glu	约 10 min
	Pro	约 7 min
	Leu、Phe、Asp、Lys、Arg	2~3 min

2. 某些特殊的氨基酸序列被用做降解信号

例如，一种名为 PEST 的序列存在于一段大概由 8 个氨基酸残基组成的肽段中，该肽段富含 P、E、S 和 T 四种氨基酸残基。酵母的转录因子 GCN4p 含有 PEST 序列，正常的 GCN4p 的半衰期约为 5 min，但如果去除 PEST 序列，半衰期就提高到 50 min。再如，一些蛋白质在 N 端含有“破坏盒”(destruction box)序列，该序列也能被一种特殊的泛素连接酶识别。

3. 信号可能隐藏在疏水核心之中

这可能就是部分折叠、变性或者异常的突变蛋白易被水解的原因。当这些蛋白质处于天然状态时，信号隐藏在疏水核心，难以被识别。一旦信号由于某种原因被暴露出来，就容易被识别而打上泛素标记。由于分子伴侣能够与疏水区域结构结合，因而能够阻止泛素的结合。

某些蛋白质的泛酰化信号可能暂时处于隐蔽状态，一方面信号由于其他蛋白质的结合而被隐藏，另一方面磷酸化修饰可影响到信号的可得性。

近年来发现的去泛酰化酶为泛酰化信号的利用披上了一层更为神秘的面纱。这一类酶的功能似乎是将已被泛素标记的靶蛋白进行去泛酰化，在它们与蛋白酶体结合之前被“抢救”过来，以避免它们的降解，从而让细胞能对一种蛋白质的量进行更为精细的调节。于是，一种蛋白质的降解除了需要特殊的泛酰化信号以外，还需要“逃脱”去泛酰化酶的“抢救”。

Quiz12 以前在哪里见过单纯由蛋白质组成的圆柱体复合体？其结构、功能与蛋白酶体有何不同？

(四) 蛋白酶体的结构与功能

蛋白酶体是由多种蛋白质组成的中空的圆柱状复合体(图 39-25)。真核生物蛋白酶体以 20S 蛋白酶体作为核心颗粒(core particle，CP)，两端各有 1 个 19S 的帽状调节颗粒(regulatory particle，RP)。其中 CP 由 14 个不同的亚基组成(每一个有 2 个拷贝)，每 7 个一组组装成环，4 个环就像 4 个面包圈一样垛叠在一起，上下两个环由 7 种 α 亚基组成，内部两个环由 7 种 β 亚基组成，蛋白酶活性中心被隔离在内部；RP 有 2 个，各位于 CP 的一端，分为底座和顶盖，至少由 17 个不同的亚基组成，其中有 6 个具有 ATP 酶活性，某些亚基含有泛素识别位点。RP 的功能首先是识别并催化泛酰化蛋白质的去折叠，然后再将它们送往位于 CP 内的蛋白酶活性中心。古菌蛋白酶体的 CP 只有 1 种 α 亚基和 1 种 β 亚基组成。

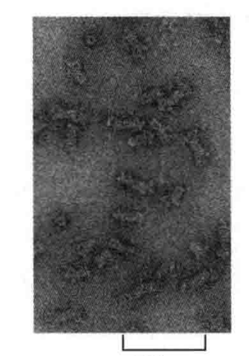




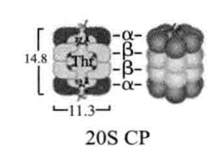
图 39-25 蛋白酶体的电镜结构

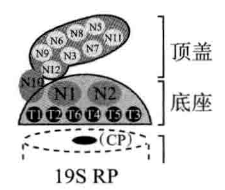
蛋白酶体具有 3 种不同的蛋白酶活性，各活性中心位于 3 个不同的 β 亚基上，其中 Thr 残基在催化中起关键作用。

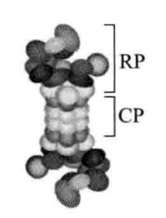
泛素标记的靶蛋白进入蛋白酶体降解的大致步骤是(图 39-26)：①RP 以依赖于 ATP 的方式，识别并结合与 Lys48 连接达到临界长度(4 个或更多的泛素分子)的多聚泛酰化蛋白质；②RP 的类分子伴侣活性对靶蛋白进行去折叠，这一步需要 ATP 的水解；③去折叠的靶蛋白通过入口进入 CP 活性中心被切成若干小肽；④带有靶蛋白残体的多聚泛素链离开蛋白酶体，在异肽酶 T(isopeptidase T)催化下，各个泛素分子得以释放。同时，被切成的小肽也从酶体中释放出来，在胞内的其他蛋白酶作用下进一步水解。

细菌既没有泛素，也缺乏与蛋白酶体相关的蛋白质，但细胞内也有依赖于 ATP 的蛋白酶降解系统。这些降解系统可有选择地降解细胞内一些异常的蛋白质，或者通过改变一种酶的浓度而调节它的活性。

e39-4 细菌内的蛋白质降解系统







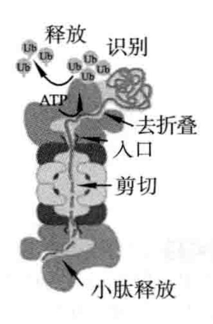


图 39-26 靶蛋白进入蛋白酶体的水解过程(单位:nm)

科学故事——"N 端规则"的发现

细胞内的蛋白质从合成出来到被分解之间,往往有一定的半衰期,而各种蛋白质的半衰期不尽相同。有的蛋白质一旦被翻译出来就很稳定,其半衰期较长,甚至比合成它们的细胞还长;有的蛋白质则非常不稳定,几乎是一合成出来就很快被分解掉了。一般说来,组织细胞必要的结构蛋白半衰期都比较长,而对于负责细胞特殊应变的调节蛋白,其半衰期往往就很短。但科学家曾一直不清楚细胞是使用何种机制来决定蛋白质半衰期的。直到 20 世纪 80 年代,MIT 的 Alexander Varshavsky 终于摸索到了一种规律。这种规律的核心内容是,蛋白质在细胞内的分解速率一般由 N 端氨基酸的性质来决定,因此被称作 N 端规则。

Varshavsky 起初只是对泛素的功能有兴趣,有一次他获悉,东京大学的 Yamada 得到一种有趣的温度敏感型细胞突变株,这种变种细胞在高于允许温度的条件下不会分裂。而 Yamada 也发现,这种突变细胞主要的问题在于它无法将泛素连接到别的蛋白质上,原因是负责连接反应的酶发生了变异,在允许温度之上就失去了活性。Varshavsky 就想到利用这种突变细胞来探讨泛素和蛋白质分解的问题。当时,许多科学家都认为,泛素是决定蛋白质半衰期的一个重要因素。根据这个理论,泛素会和那些要被消化的蛋白质结合,而促进它们的分解。如果这个理论是对的,Yamada 发现的突变细胞在允许温度之上就无法将泛素接到蛋白质上,在这样的条件下突变细胞内的蛋白质就无法被分解。

Varshavsky 做了一个简单的实验,将细胞放在允许温度之上培养,然后直接观察细胞内一些半衰期很短的蛋白质。结果发现,半衰期短的蛋白质在允许温度之上都变得十分稳定。这是第一个直接的证据显示,如果泛素不加到蛋白质上,蛋白质就不会被分解。泛素通常是接在蛋白质氨基酸侧链的游离氨基上。但利用基因工程的方法,可以把泛素接到一种蛋白质的 N 端,制造出单链的融合蛋白,然后通过直接比较这种蛋白质在接上泛素前后被分解的情形,就可以了解泛素在蛋白质分解过程中所扮演的角色。

Varshavsky 挑选了一种细菌蛋白——β- 半乳糖苷酶作为研究的对象。他把泛素的基因接到该酶基因的 5′端,然后将这个重组融合基因分别送进大肠杆菌和酵母菌。在大肠杆菌中,重组基因如期翻译出了融合蛋白——泛素 -β- 半乳糖苷酶,这表明重组基因的结构和表达都没有问题。但在酵母菌中,这个重组基因的表达却出现了寻常,它只能得到单独的β- 半乳糖苷酶。究其原因,原来是在酵母菌里,一旦泛素 -β- 半乳糖苷酶被表达,N 端的泛素就立刻被切掉。而在细菌里没有这种机制存在,因此泛素可以稳定地结合在β- 半乳糖苷酶的 N 端。为了要克服这个困难,Varshavsky 就想,是否可以通过变换β- 半乳糖苷酶在 N 端的氨基酸,来得到一种比较稳定的泛素 -β- 半乳糖苷酶融合蛋白。

Varshavsky 利用定点突变的方法,把β- 半乳糖苷酶 N 端的氨基酸做了 16 种不同的替换,结果发现只有 Pro 在 N 端时,泛素才不会立即被切除。但因此同时,β- 半乳糖苷酶本身的稳定性立刻降低了许多,其在细胞内的半衰期从二十几个小时一下降低到只有 7 min。

Varshavsky 根据上述结果得出了几个重要的结论:①在真核细胞中,切除泛素的酶不太介意β- 半乳糖苷酶的 N 端氨基酸是什么(Pro 除外);②当泛素接在蛋白质的 N 端时,也会加速蛋白质在细胞内的分解;③因为所有蛋白质合成都利用 AUG 作为起始密码子,所以 N 端最初的氨基酸都是 Met,但大多数蛋白质在后加工的时候会将其切除,而导致成熟的蛋白质在 N 端的氨基酸不一定还是 Met。这让分子生物学家很难合成一种 N 端氨基酸可任意选定的蛋白质。但利用真核细胞会准确切除融合蛋白 N 端泛素的特性,就可以先改变蛋白质基因上决定 N 端氨基酸的碱基序列,再将编码泛素的碱基序列接上。这个融合基因在真核细胞表达后,泛素会立刻被切除,而直接得到含有特定 N 端氨基酸的蛋白质了。

到此为止,Varshavsky 并没有达到最初的实验目的,但他在酵母中得到了 15 种带不同 N 端氨基酸的

β-半乳糖苷酶。接着他就顺便测定一下这些β-半乳糖苷酶在细胞内的稳定性。他首先将带有各种不同β-半乳糖苷酶基因的质粒导入到酵母中，然后将这些酵母菌放进含有[^{35}S]-Met 的培养基中生长 1~5 min，在其间合成的β-半乳糖苷酶因含有放射性[^{35}S]-Met 而具放射性，然后将酵母菌转移到没有放射性 Met 的培养基中，隔一段时间再检查酵母菌内β-半乳糖苷酶的放射性。如果β-半乳糖苷酶非常稳定，那标记过的β-半乳糖苷酶就一直能被检测到；相反，如果β-半乳糖苷酶不稳定，那β-半乳糖苷酶上的标记就会很快消失。

然而，结果大大出乎他的意料，这些仅在 1 个 N 端氨基酸有别的β-半乳糖苷酶，在酵母中半衰期的差异从 20 多个小时到 2 min。由此看来，N 端氨基酸的性质似乎直接影响到蛋白质在细胞内的稳定性。为了要进一步确定这个 N 端规则的准确性，Varshavsky 在计算机数据库中，找了 208 种位于细胞质中的稳定蛋白，发现这些蛋白的 N 端氨基酸无一例外地全属于稳定的一组。

有趣的是，他检查了 94 种分泌性的毒蛋白，发现超过 80% 的毒蛋白，其 N 端氨基酸属于非常不稳定的一组，而其他一些分泌性蛋白也有这种倾向。因此，N 端规则主要适用于最终定位于细胞质中的蛋白质。那为什么分泌性蛋白分泌出细胞外多半相当稳定，但却有不稳定的 N 端氨基酸呢？对此的解释是，如果它们在制造过程中，不小心漏进细胞质内，就会很快地被分解掉。这个解释十分合理地说明细胞在制造分泌性毒蛋白的时候，为什么不会伤害到自己，原因是毒蛋白的合成主要局限在内质网里，在那里 N 端规则不适用，若有外泄则立刻被分解。

由此可见，Varshavsky 发现 N 端规则的过程可谓是“有心栽花花不开，无意插柳柳成荫”，这为科学进展的偶发性，提供了又一个很好的例子。

e39-5 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Translation_(biology)（维基百科有关翻译的内容）
2. http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/T/Translation.html（一个免费的生物学网站，有关翻译的内容）
3. http://www.rpi.edu/dept/bcbp/molbiochem/MBWeb/mb2/part1/translate.htm（美国伦斯勒理工学院的一个分子生物化学课程网站，有关蛋白质结构的概述）
4. http://themedicalbiochemistrypage.org/protein-synthesis.html（完全免费的医学生物化学课程网站，有关翻译的内容）

参考文献：

1. Dinman J D. The eukaryotic ribosome: current status and challenges. *J Biol Chem*, 284(18): 11 761–11 765, 2009.
2. Thomas A, Steitz A. Structural understanding of the dynamic ribosome machine. *Nat Rev Mol Cell Biol*, 9: 243–253, 2008.
3. Florentz C, et al. Universal rules and idiosyncratic features in tRNA identity. *Nucleic Acid Res*, 26: 5 017–5 035, 1998.
4. Cech T R. The ribosome is a ribozyme. *Science*, 289: 878–879, 2000.
5. Baumeister W, et al. The 26S proteasome: a molecular machine designed for controlled proteolysis. *Annu Rev Biochem*, 68: 1 015–1 057, 1999.

第四十章　蛋白质的翻译后加工及其定向和分拣

不论是原核生物还是真核生物，绝大多数蛋白质在合成结束以后，还必须经历后加工才能成为有活性、有功能的形式，有的后加工反应甚至在肽链还在合成的时候就已经展开。此外，一种蛋白质在细胞中的分布不是随机的，而是被限制在特定的亚细胞空间，只有通过正确的定向（targeting）和分拣（sorting），才能到达正确的地点去行使正确的功能。

本章将主要介绍细胞内各种形式的翻译后加工反应，以及蛋白质的共翻译或翻译后的定向与分拣机制。

第一节　翻译后加工

翻译后加工概括起来主要有：多肽链的剪切和剪接、N端添加氨基酸、个别氨基酸的修饰、添加辅助因子和寡聚化。有时，蛋白质的折叠（参见第二章“蛋白质的结构”）也可以视为一种特殊形式的后加工。

1. 多肽链的剪切

Quiz1 试至少举出两个多聚蛋白的例子。

多肽链的剪切是在特殊的蛋白酶催化下进行的。许多蛋白质必须经历剪切反应，丢掉一些氨基酸序列以后才能成为有功能的蛋白质。如信号序列的切除和酶原的水解激活。某些蛋白质在翻译后以多聚蛋白质的形式存在，或者与其他蛋白质融合在一起，也需要通过剪切才能得以释放。

2. N端添加氨基酸

这是在氨酰-tRNA：蛋白质转移酶的催化下进行的，添加氨基酸的场所并非核糖体。例如，在某些真核细胞中，有一种精氨酰-tRNA：蛋白质转移酶，能催化精氨酰-tRNA分子中的Arg转移到靶蛋白分子的N端。N端添加氨基酸可改变一种蛋白质N端的性质，从而影响到它的稳定性（参见第三十九章“蛋白质的生物合成及其在细胞内的降解”）。

3. 蛋白质的剪接

蛋白质剪接是指将一条多肽链内部的一段称为内含肽（intein）的序列切除，同时将两侧称为外显肽（extein）的序列连接起来的翻译后加工方式（图40-1）。根据2011年4月份的数据，已在36种不同类型的蛋白质中发现523个内含肽。这些内含肽长度在100~800个氨基酸残基之间，遍布古菌、细菌和低等真核生物中，均位于单细胞生物。含有内含肽的蛋白质覆盖了代谢酶、DNA聚合酶、RNA聚合酶、蛋白酶、核苷酸还原酶和V型H^+-ATP酶等（参见 http://www.neb.com/neb/inteins.html）。

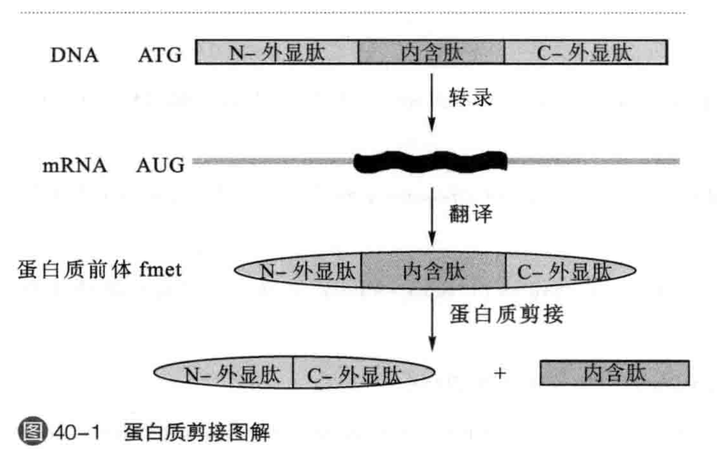


图40-1　蛋白质剪接图解

有一些蛋白质的内含肽是断裂（split intein）的，即两条肽链各含有一部分内含肽序列，需要通过反式剪接，才能将断裂的内含肽两侧的外显肽连接起来，形成有功能的蛋白质。例如，来源于集胞藻

(*Synechocystis* sp.)的 DNA 聚合酶 α 亚基,它的基因在基因组之中断裂成两个部分。这两个部分各自独立地转录和翻译,但翻译后的两条多肽前体分别在 C 端和 N 端含有断裂的内含肽。只有通过反式剪接,才能将位于两条肽链上的外显肽连接起来,形成一个连续的有功能的 α 亚基。

Quiz2 如何设计一个实验,证明蛋白质的剪接属于自催化反应?

就剪接的机制来看,蛋白质剪接经历一种分支蛋白中间体,是一种自催化反应,而释放出来的内含肽通常具有核酸内切酶的活性。

框 40–1　生化探寻——内含肽的"内涵"

蛋白质在剪接过程中,释放出来的内含肽通常具有位点特异性 DNA 酶的活性。这种酶属于归巢内切核酸酶,具有长的识别序列(14~40 bp),通过锁定和切开无 I 组内含子和内含肽的基因,而引发宿主细胞内的 DNA 修复系统以含有内含子和内含肽的基因为模板,从而将编码内含子和内含肽的序列拷贝到新的靶点,而促进 I 组内含子和内含肽的移动和传播。

这种酶活性的功能是保护内含肽的存在。如果一个宿主基因中间编码内含肽的序列发生缺失突变,宿主细胞不会受到伤害,但是内含肽就丧失了。然而,倘若内含肽的 DNA 酶活性在一个宿主基因的中间切开 DNA,这可能会杀死任何缺失了无用的内含肽的细胞。只有保留内含肽的细胞才能生存下来。这种切割通常发生在插入内含肽的地方。含有内含肽的基因不会被切,原因是识别位点由于内含肽的插入而被破坏了。内含肽对细胞的生存不是必需的,编码内含肽的 DNA 序列因此被称为一种自私的 DNA(只有利于自己的生存,而对宿主没有什么好处)。

4. 个别氨基酸的修饰

氨基酸的修饰包括对 N 端或 C 端的修饰以及对氨基酸侧链的修饰。主要的修饰方式包括磷酸化、亚硝基化、糖基化、瓜氨酸化、脂酰化、异戊二烯化、羟基化、泛酰化、ADP– 核糖基化、甲基化、酰胺化、乙酰化、甲酰化、γ– 羧基化、碘基化、焦谷氨酰化、硫酸化、GPI 附着、AMP 化或 UMP 化、小泛素相关修饰物修饰。氨基酸残基的修饰不仅能改变蛋白质的某些理化性质,还能调节许多酶或蛋白质的活性。

e40–1 蛋白质分子中各种形式的氨基酸残基修饰

5. 添加辅助因子

许多蛋白质或酶在合成以后,必须与相应的辅助因子结合才有活性,比如绝大多数羧化酶需要生物素作为辅助因子,细胞色素和血红素蛋白则需要血红素辅基。

6. 寡聚蛋白质的寡聚化

这是具有四级结构的蛋白质才有的后加工方式。蛋白质的寡聚化是一种自组装过程,通常也涉及分子伴侣。分子伴侣在其中所起的作用是保护亚基的疏水表面,直到各亚基有机会接触并形成寡聚体。

第二节　蛋白质翻译后的定向转运与分拣

每一个活细胞每时每刻都在合成多种蛋白质,合成的场所主要是细胞质,但不同的蛋白质在合成以后,经过定向与分拣,最后都到达目的地。比如,组蛋白进入细胞核,而胰岛素被分泌到胞外。那么细胞中的蛋白质究竟是怎样各得其所的呢?

如果你对信件的分拣和投递过程有所了解,也就不难理解蛋白质的定向和分拣的基本原理。信封上的邮政编码很重要,由它决定一封信的最终去向,但邮局里那台能够自动识别邮政编码的机器同样重要,没有它的识别,邮政编码也就没有任何意义。蛋白质分子上有没有类似"邮政编码"的信号呢? 要是有,那么用来识别这种"分子邮政编码"的机器又如何发挥作用呢?

一、信号假说

对于新生的蛋白质有没有指导其定位和分拣的“分子邮政编码”，Blobel 和 Dobberstein 选择了免疫球蛋白为对象，进行了系统的研究，并于 1975 年提出了信号假说（signal hypothesis）。信号假说的主要内容是：各种蛋白质在细胞中的最终定位是由蛋白质本身所具有的特定氨基酸序列决定的。这些特殊的氨基酸序列起着一种信号向导的作用，因此被称为信号序列。信号序列能够被细胞中的特殊成分识别，由此启动定向和分拣的过程。一个蛋白质分子如果缺乏任何一种信号序列，就会留在细胞质基质，如参与糖酵解和磷酸戊糖途径的所有酶。整个定向和分拣的过程可分为识别、移位和成熟。

到现在为止，已发现多种形式的信号序列，这些信号序列不仅参与蛋白质的定向和分拣，还可能具有其他的功能（表 40–1）。此外，有些蛋白质天生没有任何信号序列，而有些蛋白质则有多重信号序列（表 40–2）。

► 表 40–1 常见的信号肽序列

信号类型	基本特征
SRP 信号	位于进入内质网、高尔基体、溶酶体、细胞膜和胞外的蛋白质的 N 端，一般由 20 个氨基酸残基组成，多数为疏水氨基酸，能被 SRP 识别
内质网滞留信号	位于 C 端，一致序列为 –KDEL
高尔基体滞留信号	一段由 20 个疏水氨基酸组成的肽段，两侧是带正电荷的碱性氨基酸
溶酶体分拣信号（信号斑）	由一级结构上并不相邻的若干氨基酸残基构成的三维信号，其中带正电荷的氨基酸残基十分重要
线粒体分拣信号（导肽）	位于 N 端，由 15~70 个氨基酸残基组成，内有一组带正电荷的碱性氨基酸残基，几乎无酸性氨基酸，其中穿插一些亲水氨基酸残基，在水溶液中很少形成二级结构，但是一旦插入线粒体膜则会自动形成两亲螺旋
叶绿体分拣信号（输送肽）	位于大多数叶绿体蛋白的 N 端，由 50 个左右的氨基酸残基组成，一般富含 Ser、Thr 和一些小的亲水氨基酸残基，很少有酸性氨基酸
核定位信号（NLS）	位于肽链内部，由 4~8 个富含碱性的氨基酸组成，还常有 1 个或几个 Pro
核输出信号（NES）	位于可以进入又可以离开细胞核的蛋白质，一致序列为 LXXXLXXLXL，L 为疏水氨基酸，通常是 Leu，X 代表任何氨基酸
过氧化物酶体分拣信号（PTS）	有两种，一种位于肽链的 C 端（PTS1），其一致序列为 –S/A–K/R–L/M，其中以 –SKL 最常见，另一种位于 N 端（PTS2），是一段九肽序列

► 表 40–2 不同的蛋白质所具有的信号序列的数目和种类

最后的目的地	信号数目和种类	最后的目的地	信号数目和种类
细胞质基质	无	质膜	SRP 信号和停止转移信号
内质网	SRP 信号和内质网滞留信号	线粒体基质	线粒体分拣信号（导肽）
高尔基体	SRP 信号和高尔基体滞留信号	叶绿体基质	叶绿体分拣信号（输送肽）
溶酶体	SRP 信号和溶酶体分拣信号（信号斑）	细胞核	核定位信号（NLS）
分泌小泡	SRP 信号	过氧化物酶体	过氧化物酶体分拣信号（PTS）

进一步研究表明，蛋白质在细胞内的定向和分拣有两条途径：一条途径与翻译同时发生，即共翻译途径；另一条途径在翻译完成以后才进行，即翻译后途径。

二、蛋白质的共翻译定向

定位于内质网、高尔基体、溶酶体、分泌小泡、细胞膜和胞外的蛋白质都属于共翻译定向。最早发现的信号序列就是由 Blobel 发现的 SRP 信号序列，其特征见表 40–2。它存在于内质网蛋白、高尔基体蛋白、溶酶体蛋白、细胞膜蛋白和分泌蛋白的 N 端。由于信号肽位于 N 端，因此总是最先被翻译，当翻译达到约 80 个氨基酸残基的时候，信号肽序列便突出于核糖体的表面，被 SRP 识别，由此启动定向和分拣过程。

SRP 是一种特殊的核糖核酸蛋白颗粒，在细菌和叶绿体内也有发现。哺乳动物的 SRP 由一种

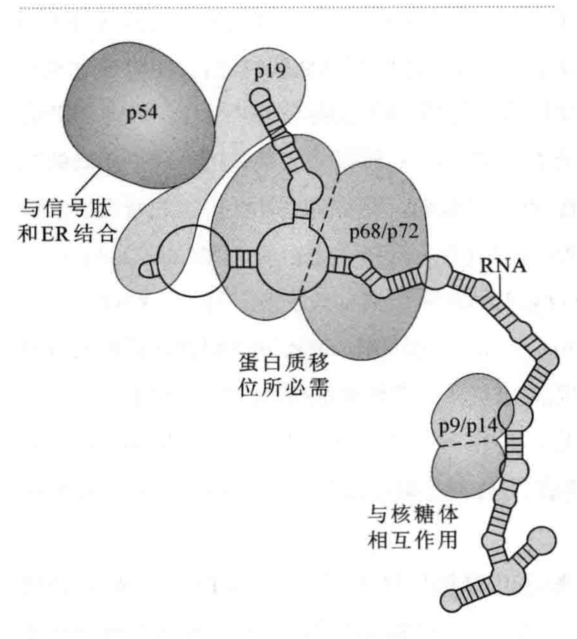


图 40-2 SRP 的结构与功能(Lodish 等,2007)

7SL RNA 以及六种多肽(p9、p14、p19、p54、p68 和 p72)组成(图 40-2)。7SL RNA 可被切割成两个部分:小片段结合 p9/p14,能阻滞翻译;大片段结合 p68/p72 和 p54/p19,引发移位,其中 p54 分别通过 p54M 和 p54G 位点识别信号序列和结合 GTP。叶绿体和大肠杆菌的 SRP,则是由一分子的 4.5S RNA 和一种大小为 4.8×10^4 的多肽组成。

SRP 作为一种接头分子,含有一个大的疏水口袋。此口袋为信号序列结合部位,内部以 Met 居多。Met 含有灵活的侧链基团,这对于 SRP 与信号序列建立疏水作用再合适不过了。如图 40-3 所示,一旦新生的肽链从大亚基上的肽链离开通道内探出头来,SRP 即与暴露出来的信号序列结合。SRP 的结合使翻译暂停,从而让核糖体在翻译恢复之前有足够的时间与内质网膜结合。然而,核糖体并不能直接与内质网膜结合。在它与内质网结合之前,需要 SRP 与内质网膜上的 SRP 受体结合。一旦 SRP 结合到它自己的受体上,核糖体就能与内质网膜上相应的受体以最佳的角度结合。由于 SRP 受体和 SRP 都是 G 蛋白,在它们相互结合的时候,各自的 GTP 酶活性被激活。于是,与它们结合的 GTP 发生水解。GTP 的水解导致 SRP 与信号肽解离。一旦 SRP 解离,肽链即恢复合成。随后,信号肽通过膜上的移位子开始向内质网腔转移,新生肽链则尾随信号肽继续延伸。

Quiz3 如果没有 SRP,真核细胞还有没有糙面内质网?

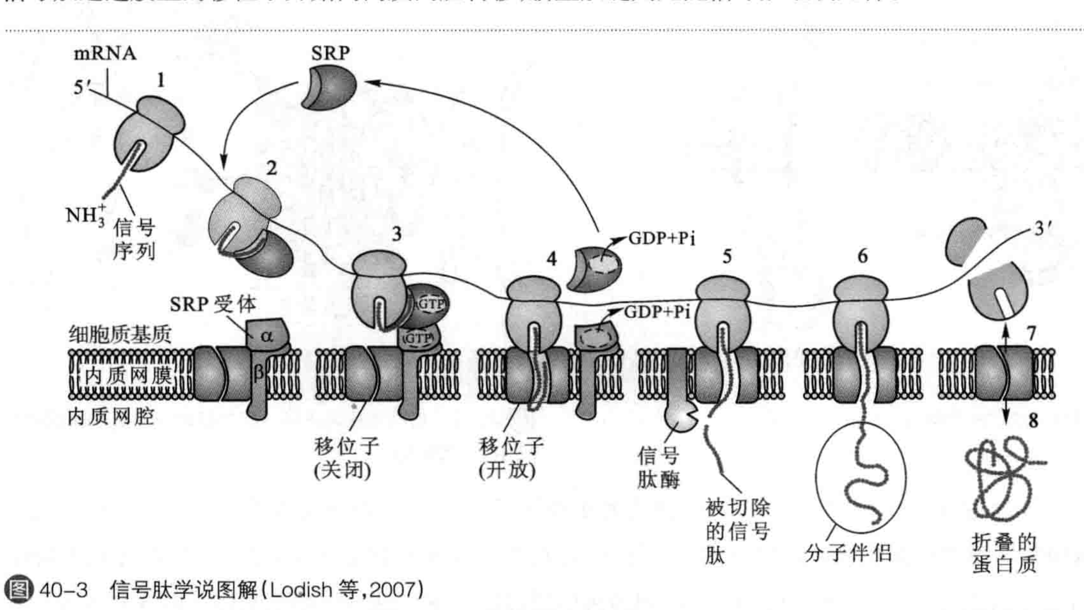


图 40-3 信号肽学说图解(Lodish 等,2007)

调节蛋白质通过内质网膜转运的移位子称为 Sec61 复合体。它含有水溶性的孔道,允许正在延伸的多肽链发生跨膜转移。在核糖体与 Sel61 复合体结合的时候,核糖体大亚基上的肽链离开通道与 Sel61 复合体中央孔无缝对接,核糖体内的空间与内质网腔连为一体,这样的结构使得从核糖体上伸出来的多肽链能通过 Sec61 的中央孔进行移位,同时还能有效防止内质网腔内 Ca^{2+} 和其他物质的泄漏。

正是信号序列触发了中央孔的开启,并引导蛋白质通过内质网膜。蛋白质像线一样以环的形式连续地通过膜。信号肽在穿过膜后,被“守候”在内质网膜内侧的信号肽酶(signal peptidase)切下,新生肽链则继续延伸直至终止。

如果不是膜蛋白,那一旦 C 端通过移位子,蛋白质就会被释放到内质网腔。释放到内质网腔的每一种蛋白质都需要进行折叠和 / 或组装成寡聚体,而折叠和组装需要腔内的分子伴侣的帮助。内质网

腔内最丰富的分子伴侣是 BiP。BiP 主要参与糖蛋白以外的蛋白质的折叠。其作用方式与其他分子伴侣一样，即与新生肽链上的疏水区结合，阻止多肽链在没有折叠之前与其他肽链聚合，同时通过水解 ATP 驱动多肽链的折叠。多肽链如果折叠正确，疏水区得以包埋，就不会再与 BiP 结合。否则，BiP 会再次与其结合，重新启动折叠。含有二硫键的蛋白质在折叠的时候，伴随着二硫键的形成。正确的二硫键的形成需要内质网上的蛋白质二硫化物异构酶（PDI）的帮助。那些折叠频繁出错的蛋白质会激活内质网内的多种形式的质量控制系统。其中一种控制系统称为无折叠蛋白质应答（unfolded protein response，UPR）。该系统能对错误折叠的蛋白质进行检测，然后再激活过载反应。另外一种就是与内质网关联的蛋白质降解途径（ER-associated degradation pathway，ERAD）。该系统识别错误折叠或者没有组装好的蛋白质，并将它们逆向转移到细胞质基质，在细胞质基质再被送到蛋白酶体降解。

如果是膜蛋白，那肽链中就还可能含有另外一种信号序列，即停止转移序列（stop-transfer sequence）（图 40–4），该序列约由 20 个疏水氨基酸组成，其作用是阻止肽链的进一步移位，以让肽链在 N 端信号序列被切除以后仍然能锚定在膜上。

对于绝大多数通过共翻译途径转移的蛋白质来说，内质网仅仅是其定向转运的第一站，其他的站点还包括高尔基体、溶酶体、细胞膜和细胞外等。一般而言，相邻站点之间通过小泡转运进行联系（图 40–5）。那些以内质网为"永久性居留地"的可溶性蛋白质（如 BiP）在 C 端还具有内质网滞留信号（ER retention signal）——KDEL 序列。

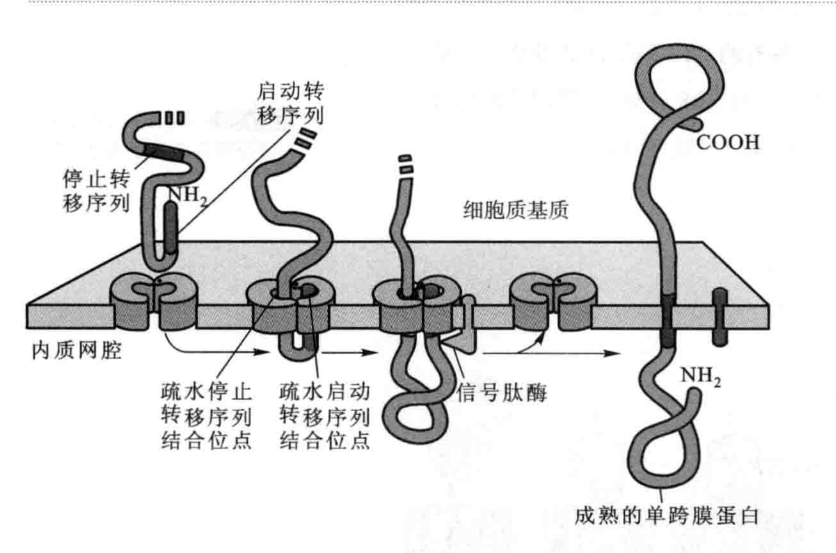


图 40–4　停止转移序列介导的单跨膜蛋白的定向转移（Alberts 等，2007）

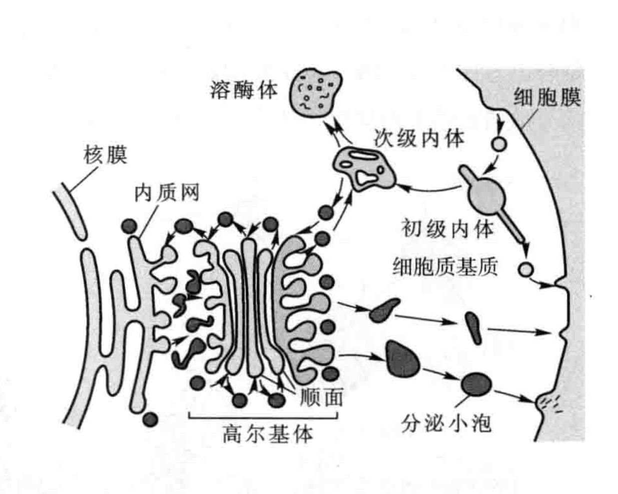


图 40–5　蛋白质从内质网→高尔基体→溶酶体或细胞膜（外）的小泡转运

实际上，蛋白质每进入一个站点，都需要做出抉择：是进入下一站还是以此为终点站？在 C 端含有 KDEL 序列的可溶性蛋白就以内质网为终点站，在 C 端含有 KKXX（X 为任意氨基酸）序列的跨膜蛋白也是如此，而不含 KDEL 或 KKXX 序列的蛋白质即进入下一站——高尔基体。但含有 KDEL 或 KKXX 序列的蛋白质并不一定就不会被误送到高尔基体。然而，误入高尔基体的内质网蛋白由于含有 KDEL 序列或 KKXX 序列，会在高尔基体与特定的受体蛋白结合，然后被选择性打包到特定转运小泡，又返回到内质网。

Quiz4 你认为哺乳动物细胞的 HMG-CoA 还原酶含有多少种不同的信号序列？

进入高尔基体的蛋白质又要面临新的选择。以溶酶体为目的地的蛋白质带有某种特殊的被称为信号斑（signal patch）的三维信号，这种信号使得它们能够发生一种特殊的化学修饰——在其共价相连的 N 寡糖链上再引入一个或多个 6- 磷酸甘露糖（M6P）。在它们发生 M6P 修饰以后，位于高尔基体反面（*trans* side）膜上的 M6P 受体蛋白能识别并结合它们。随后，这个区域的高尔基体膜在网格蛋白的帮助下，以出芽的形式被打包成分拣小泡，小泡的形成需要消耗 ATP。随后，小泡与次级内体（late endosome）融合，并最终成熟为溶酶体。而分泌蛋白和细胞膜蛋白由于无 M6P，因此会在高尔基体另

外的区域，同样以出芽的方式形成分泌小泡。分泌小泡可与细胞膜融合，释放可溶性的分泌蛋白，而整合在小泡膜上的蛋白则变成了细胞膜蛋白。

三、蛋白质的翻译后定向

定位于线粒体、质体、细胞核和过氧化物酶体的蛋白质（线粒体和叶绿体本身合成的蛋白质除外），基本上都属于翻译后定向。有些蛋白质具有双重定位信号，其中有一种定位信号比较隐蔽，这些蛋白质可定位到不同的细胞器。

（一）蛋白质定位于线粒体

线粒体蛋白质组大概含有 1 500 种不同的蛋白质，但只有 10 多种由线粒体基因组编码，绝大多数蛋白质由核基因编码。这些由核基因编码的蛋白质在细胞质游离的核糖体上翻译，经过翻译后定向和分拣进入线粒体。构成它们的肽链含有线粒体分拣信号，其中最重要的信号位于 N 端，这种信号序列有时被称为导肽（leader peptide）。导肽的序列特征见表 40-2。这种信号序列不会被 SRP 识别，在水溶液中很少形成二级结构，但是一旦插入线粒体膜就会自动形成两亲螺旋。

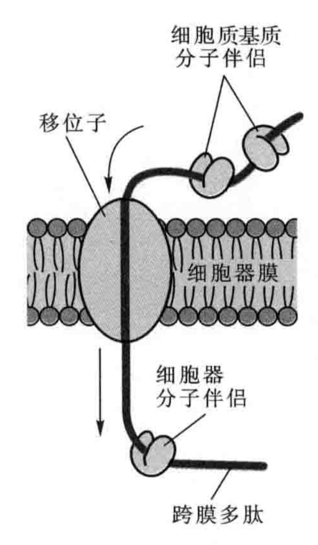


图 40-6　分子伴侣在翻译后的蛋白质定向转移中的作用

为了便于后面的跨膜转运，细胞质中的分子伴侣与蛋白质前体结合，以防止其提前折叠或相互间“非法”聚焦，维持蛋白质处于一种容易跨膜的伸展状态。在线粒体基质一侧，也有分子伴侣与正在入内的肽链结合（图 40-6）。

然而，线粒体的亚空间包括基质、膜间隙、内膜和外膜。进入线粒体的蛋白质又如何被分拣到最后的亚空间呢？一般认为，不同的蛋白质还可能带有次级信号序列，正是这些次级信号序列决定了一种线粒体蛋白在线粒体最终的位置。

1. 线粒体基质蛋白的定向转运

如图 40-7 所示，在细胞质基质，分子伴侣 HSP70 与正在游离的核糖体上合成的肽链结合，通过水解 ATP 以阻止其提前折叠。以线粒体为目的地的蛋白质多数在 N 端含有导肽序列，这种信号序列使其能与横跨外膜的转运蛋白（transport across the outer membrane，TOM）复合体上的受体蛋白结合，以方便蛋白质与移位子之间的相互作用。为了让蛋白质能够顺利跨膜，细胞质基质的 HSP70 必须与结合的肽链解离，此过程需要 ATP 的水解。

TOM 复合体与导肽结合的部分由 Tom70 和 Tom37 组成，这两种蛋白质随后将蛋白质交给另外两种蛋白质——Tom22 和 Tom20，而 Tom22 和 Tom20 再将蛋白质转移到由 Tom40 组成的通道。在内膜上，则有另外一种叫横跨内膜的转运蛋白（transport across the inner membrane，TIM）复合体，由它构成蛋白质穿过内膜的通道。蛋白质进入基质只发生在内外膜的接触点，形成两亲螺旋有利于蛋白质的跨膜转运。TOM 和 TIM 通过静电作用结合在一起。Tim 23 和 Tim 17 形成移位子。蛋白质从 Tom40 插入到 Tim23 形成的通道需要呼吸链创造的电化学梯度（$\Delta\Psi$）。此外，在基质一侧，受 ATP 驱动的前导序列转位酶相关联的分子马达（presequence translocase-associated motor，PAM），也是蛋白质完成跨膜转运进入基质所必需的。PAM 的主要成分是线粒体内的 HSP70（mtHSP70），一旦肽链跨入基质，mtHSP70 就与其结合，通过水解 ATP 来阻止其折叠，并驱动它的跨膜转运。导肽在进入基质以后，受基质内信号肽酶的作用而被切下。丢掉导肽的基质蛋白进入由 mtHSP60 和 mtHSP10 形成的桶装折叠体的内部，受 ATP 水解的驱动进行折叠。

2. 线粒体非基质蛋白的定向转运

蛋白质进入基质以外的地方是从 TOM 复合体开始的，对定位到内膜的蛋白质而言，至少有三种不同的方式：①有些蛋白质像基质蛋白一样，具有导肽的前序列。这些蛋白质也需要 TOM/TIM 复合物，但由于在导肽序列之后就是一段疏水的停止转移信号序列，因此它们不需要 PAM 输入马达，最后直接留在内膜上即可。②有些蛋白质为内膜上的多次跨膜蛋白（如载体蛋白），这些蛋白质没有导肽序列，但内部含有多重线粒体输入信号。这些蛋白质起初由外膜上另外一种受体蛋白——Tom70 识别，然

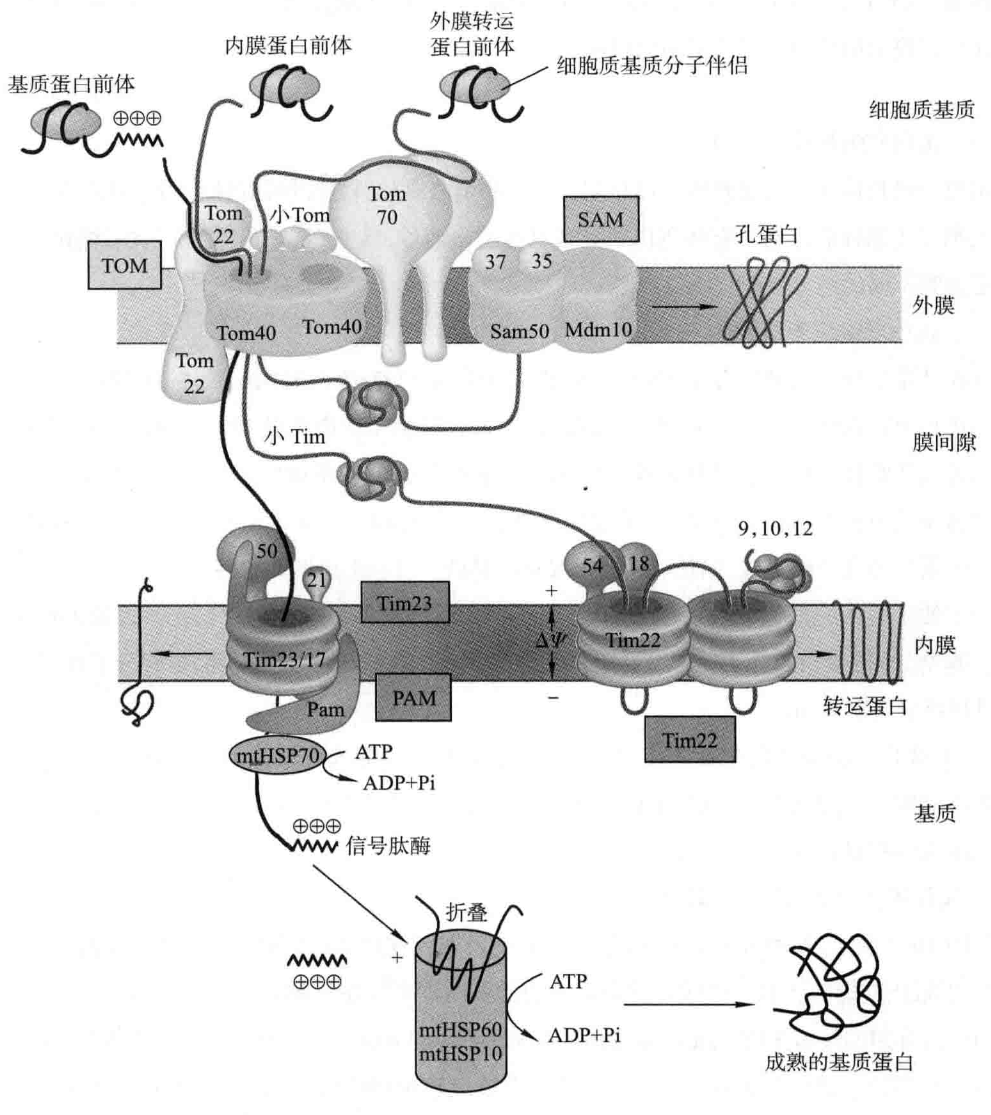


图 40-7　细胞核编码的线粒体基质蛋白的定向转移(Goodman,2008)

后再通过 Tom40 转运,在膜间隙,被一些小的可溶性 Tim(Tim9/Tim10)蛋白质复合体识别,随后,受内膜上的 Tim22 复合物的作用,以及内膜内外电化学势能的驱动,最终插入到内膜上。③还有一些蛋白质最初像基质蛋白一样,被输送到基质,但在基质被切掉导肽序列以后,会暴露出一种新的能被 Oxa1 蛋白识别的次级信号序列。在 Oxa1 蛋白的作用下,这些蛋白质也被插入到线粒体内膜上。需要注意的是,Oxa1 也参与转运那些由 mtDNA 编码的内膜蛋白。

对外膜蛋白而言,主要有两种情况:具有简单拓扑学结构(如单跨膜)的外膜蛋白只需要 TOM 复合物就可以插入到外膜上。但那些拓扑学结构较复杂的蛋白质(如具有 β 桶结构的孔蛋白和 Tom40),首先需要通过 TOM 复合体进入膜间隙,然后在膜间隙中较小的 Tim 蛋白的帮助下,通过 SAM 复合物插入到外膜。

(二) 蛋白质定位于叶绿体

与线粒体相比,叶绿体含有更多的蛋白质,而许多蛋白质是由核基因编码,例如,DNA 聚合酶、氨酰 -tRNA 合成酶、约 2/3 的核糖体蛋白、参与光反应的部分蛋白质和参与暗反应的 Rubisco 的小亚基。由核基因编码的蛋白质进入叶绿体的过程(图 40-8)与进入线粒体很相似,但至少也有三大差别:①线粒体比它少一层膜(类囊体膜)和一种可溶性腔(类囊体腔),因此在类囊体蛋白进入叶绿体基质,还需要进一步分拣和定位,才能被运送到类囊体;②在结构上,参与叶绿体膜转运的 Toc 或 Tic 与参与线粒体膜转运的 Tom 或 Tim 无同源性;③蛋白质进入叶绿体体基质不需要也没有跨内膜的质子梯度。

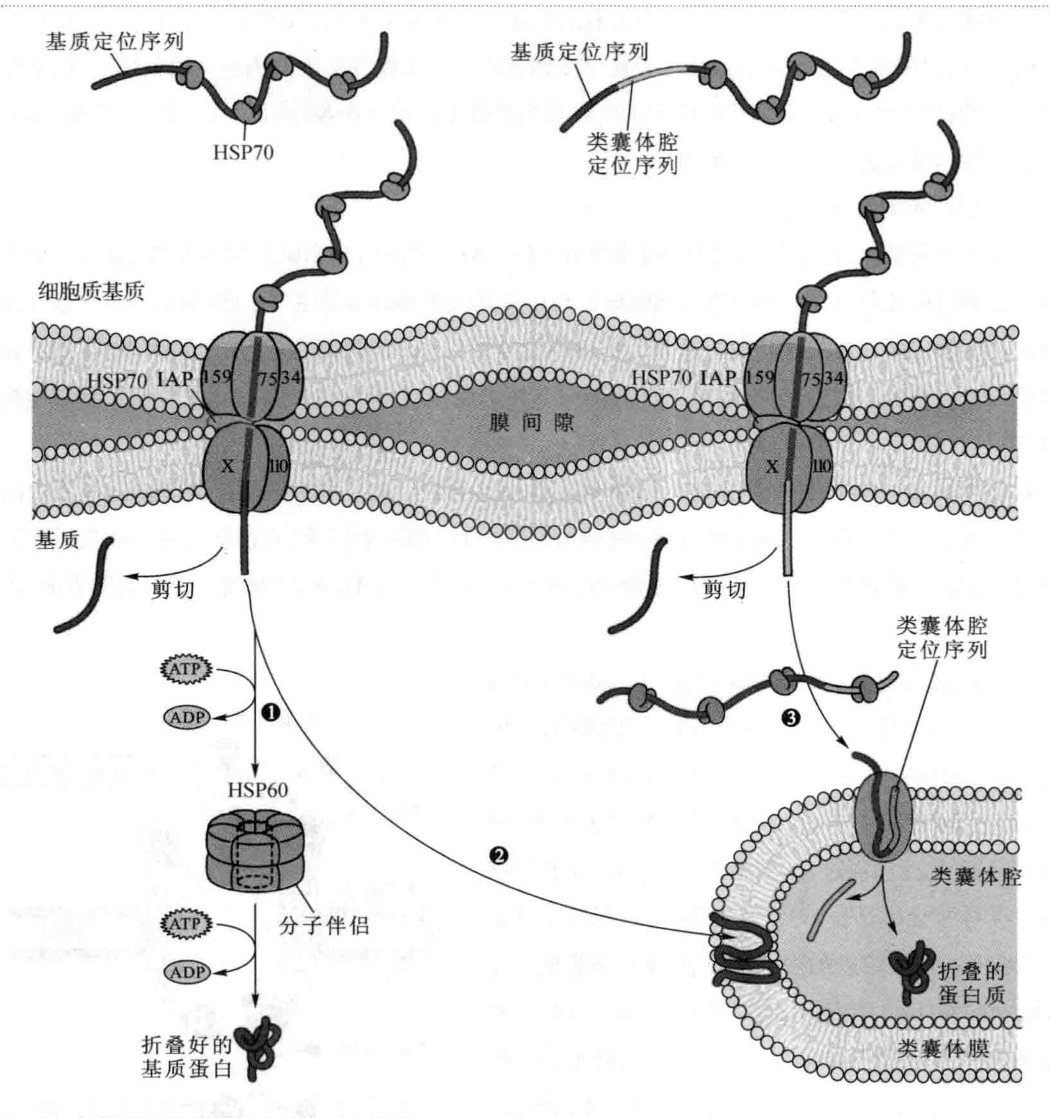


图 40-8 细胞核编码的叶绿体蛋白的定向转移

1. 定位于基质

进入叶绿体的蛋白质在细胞质基质与分子伴侣 HSP70 结合而处于伸展状态，在 N 端含有叶绿体分拣信号，这段信号序列被称为输送肽（transit peptide），其特征见表 40-1。含有输送肽序列的蛋白质通过受体介导的转运方式，进入叶绿体的基质，在跨膜转移中，需要消耗 ATP 和 GTP。它们进入基质以后，会被特殊的蛋白酶在 N 端切掉 20~25 个氨基酸残基。以基质为终点的蛋白质，要进入基质内由 HSP60 组成的桶装复合物的内部进行折叠；以类囊体为终点的蛋白质，则在新暴露出来的次级信号序列的指导下，作进一步的分拣和定向。

转运受体和移位子复合体在内外膜接触点组装，包括 3 个部分：一是横跨叶绿体外膜的转运蛋白（transport across the outer chloroplast membrane，TOC）复合体，由 1.59×10^5、7.5×10^4 和 3.4×10^4 蛋白组成，其中 1.59×10^5 和 3.4×10^4 蛋白结合 GTP；二是横跨叶绿体内膜的转运蛋白（transport across the inner chloroplast membrane，TIC）复合体；三是 HSP70 转运中间物关联蛋白（import intermediate associated protein，IAP），位于 TOC 和 TIC 之间。

2. 定位于类囊体膜和腔

定位于类囊体膜和腔的蛋白质具有较长的输送肽序列，其内部包括两种信号：第一种信号指导蛋白质通过内外膜进入叶绿体基质，第二种信号比较隐蔽，它负责指导蛋白质从基质进入类囊体。这两种信号按照一定的次序被切除。在蛋白质进入叶绿体的基质以后，第一种信号序列被切除。与此同时，第二种信号随之暴露，这种信号序列为类囊体分拣信号，负责指导蛋白质进入类囊体。

Quiz5 若是让 *GFP* 基因在叶肉细胞中表达，并最终定位到类囊体腔内，你应该在 *GFP* 上加上几种信号序列？

类囊体分拣信号序列是由一段 25 个左右的疏水氨基酸残基组成，其前端常常有一些带正电荷的氨基酸残基，以便和类囊体膜上的移位子复合体相互作用。正如许多蛋白质进入线粒体一样，某些蛋白质进入类囊体也需要消耗跨膜的质子梯度。当类囊体蛋白进入类囊体以后，信号序列会被切除，其内的分子伴侣可促进它们正确地折叠。

（三）蛋白质定位于细胞核

定位于细胞核的蛋白质在肽链的内部含有 NLS。NLS 主要由几个碱性氨基酸残基组成，负责引导蛋白质通过核孔复合物（NPC）进入细胞核。与许多其他类型的信号序列不同的是，NLS 在指导蛋白质进入细胞核以后并没有被切除。这是因为细胞在分裂的某个阶段，核膜会发生解体，如果 NLS 在细胞核蛋白进入细胞核以后被切除，就会导致细胞分裂结束以后，原来的核蛋白无法重新回到细胞核的“怀抱”。

核蛋白通过 NPC 的转运还需要一种叫 Ran 的小 G 蛋白。Ran 的功能就是调节转运核蛋白的受体复合体进行组装或解体。核蛋白从核糖体释放出来以后即发生折叠，故核蛋白是以完全折叠的状态通过核膜进入核质的。在折叠好的核蛋白分子上，NLS 经常以环的形式暴露在蛋白质的表面，以方便识别。

核蛋白进入细胞核的详细步骤是（图 40-9）：①核蛋白与 NLS 的受体——输入蛋白 α/β 二聚体结合，形成复合物；②核蛋白 - 受体复合物与 NPC 胞质环上的纤维结合；③纤维向核弯曲，移位子构象发生改变，形成亲水通道，核蛋白通过；④细胞核内的 Ran-GTP 与核蛋白 - 受体复合体结合，导致复合体解散，释放出核蛋白；⑤与 Ran-GTP 结合的输入蛋白 β，离开细胞核。与 Ran 结合的 GTP 在细胞质水解，产生的 Ran-GDP 返回细胞核，并重新转变为 Ran-GTP；⑥在核内输出蛋白的帮助下，输入蛋白 α 被运回细胞质，参与下一轮的转运。

图 40-9 由 NLS 介导的细胞核蛋白的定向转移（Lodish 等，2007）

Quiz6 如果人为地在核蛋白的 N 端加上 SRP 信号序列，你认为这种蛋白质最后的去向在哪里？

（四）蛋白质定位于过氧化物酶体

过氧化物酶体是由单层膜包被的细胞器，它所具有的所有蛋白质均由细胞核 DNA 编码。指导蛋白质进入过氧化物酶体的信号序列被称为过氧化物酶体定向序列（peroxisome targeting sequence，PTS）。PTS 一般位于肽链的 C 端（PTS1），其一致序列为 S/A-K/R-L/M，其中以 SKL 最常见，也有少数蛋白质的 PTS 是位于 N 端的九肽序列（PTS2）。

与核蛋白一样，进入过氧化物酶体的蛋白质在细胞质基质已完全折叠成有功能的形式。此外，在它们进入目的地以后，所带的 PTS 与核蛋白上的 NLS 一样也没有被切除。

e40-2 细菌的蛋白质定位

细菌虽然没有细胞器结构，但也需要将它的一些蛋白质从细胞质输送到外膜、内膜、外周腔（内外膜之间的空隙）或胞外，因此同样需要相应的分拣和定位系统，而那些需要被转运的蛋白质在 N 端也应该含有特殊的信号序列，以指导它们的跨膜转运。

科学故事——“信号学说”的发现

1999 年的诺贝尔生理学或医学奖被授予美国洛克菲勒大学的 Günter Blobel，因为他发现了蛋白质具有内在的信号以控制它们在细胞中的转运和定位（proteins have intrinsic signals that govern their transport and localization in the cell）。那么，Blobel 是如何取得这一发现的呢？

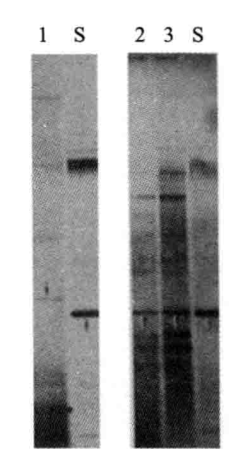
生物化学家曾一直想搞清楚：细胞内一个特定的多肽链是如何横跨生物膜，进入属于自己的亚细胞

空间或被分泌到细胞外的？20 世纪 50~60 年代的一些研究结果表明，分泌蛋白是在与膜结合的核糖体上合成并同时完成跨膜转移的，但没能解释为什么分泌蛋白需要在与膜结合的核糖体上合成而细胞质基质蛋白不需要的现象。

直到 1971 年，Blobel 和 David Sabatini 提出了一种假说，对上述现象提出了他们的解释：①在与膜结合的核糖体上翻译的 mRNA 5′端紧跟起始密码子之后含有一段特殊的密码子；②这些密码子编码特殊的氨基酸序列；③这种特殊的氨基酸序列诱发了多肽链的跨膜转移。

1975 年，Blobel 和 Dobberstein 相继获得一系列实验证据支持他们最初的假说，在此基础上，他们对原来的学说做了一些修改，并正式命名为信号学说。

骨髓瘤是一种 B 淋巴细胞瘤，它能旺盛地合成和分泌免疫球蛋白（IgG），因此是一种非常理想的研究信号假说的模型。Cesar Milstein 实验室早期的研究结果显示，与在体内正常分泌出来的 IgG 的轻链相比，IgG 轻链的 mRNA 在体外翻译出来的多肽链在 N 端多出约 20 个氨基酸残基（图 40-10）。Blobel 和 Dobberstein 认为，正是这些多出来的氨基酸序列充当信号，指导核糖体与膜的结合。为了证实这种设想，Blobel 和 Dobberstein 继续研究 IgG 轻链在与膜结合的核糖体上合成的过程，他们既重复出 Cesar Milstein 实验室的结果，又进一步发现在与微粒体膜结合的核糖体上合成的 IgG 轻链对加入的蛋白酶水解有抗性，这说明 IgG 轻链在合成中被转移到微粒体的腔中，随后充当信号的序列被微粒体上的酶切除了。根据上述结果，Blobel 和 Dobberstein 提出了更为详细和准确的信号学说。



1. 在体外游离的核糖体上翻译的产物
2. 在体外与膜结合的核糖体上翻译的产物
3. 在体外与膜结合的核糖体上翻译的产物经蛋白酶处理
S. 正常的分泌蛋白

图 40-10 骨髓瘤细胞 IgG 轻链在不同条件下被翻译后产物的电泳图谱

1980 年，Blobel 根据对细胞内其他的一些膜蛋白和细胞器蛋白跨膜转移的研究结果，认为定向并输送到不同细胞器的新生蛋白也含有内在的信号序列，这些特殊的信号序列由原来的蛋白质基因编码，它可能位于多肽链的两端，也可能在多肽链的内部。

蛋白质高度选择性的跨膜定位和分拣对于细胞，特别是对真核细胞的功能至关重要。为了维持各种不同细胞器的个性，相关的蛋白质必须准确地到达属于自己的"领地"。信号学说不仅正确地解释了分泌蛋白是如何通过内质网最后被分泌出细胞，还为探索其他需要跨膜转移的蛋白质的定向转运机理提供了正确的方向。

e40-3 本章小结填填看

推荐网址：

1. http://themedicalbiochemistrypage.org/protein-modifications.html（完全免费的医学生物化学课程网站有关翻译的内容）
2. http://en.wikipedia.org/wiki/Signal_peptide（维基百科有关信号肽的内容）
3. http://www.columbia.edu/cu/biology/courses/c2006/lectures05/lect7.05.html（美国哥伦比亚大学一个生物学网站，有关蛋白质定向和分拣的内容）
4. http://www.cryst.bbk.ac.uk/pps97/assignments/projects/dulai/signal.html（英国伦敦大学提供的各种信号肽有关的信息）

参考文献：

1. Saleh L, Perler F B. Protein splicing in cis and in trans. *Chem Rec*, 6(4): 183–193, 2006.
2. Kutay U, et al. Protein transport across the eukaryotic endoplasmic reticulum and bacterial inner membranes. *Annu Rev Biochem*, 65: 271–303, 1996.
3. Schatz G, Dobberstein B. Common principles of protein translocation across membranes. *Science*, 271: 1 519–1 525, 1996.
4. Görlich D, Mattaj I W. Nucleocytoplasmic transport. *Science*, 271: 1 513–1 518, 1996.

第四十一章　再次程序化的遗传解码和翻译暂停

绝大多数蛋白质的合成都是以一条 mRNA 作为模板，从起始密码子开始翻译，到终止密码子结束，肽链在延伸的时候，每参入一个氨基酸，核糖体只移位一个密码子。但是，也有少数蛋白质在翻译的时候，当核糖体前进到 mRNA 的某些区段的时候，却偏离上述常规，而以一些特殊的方式进行解码，这些特殊方式的解码称为再次程序化的遗传解码（reprogrammed genetic decoding）。再次程序化的遗传解码并不违背"中心法则"，而是对它的补充和扩展。当核糖体通过这些特殊的区段以后，将重新按照一般的解码方式继续翻译。

本章将重点介绍几种再次程序化的遗传解码形式，包括翻译水平的移框、通读、跳跃翻译以及含硒半胱氨酸和吡咯赖氨酸的参入。

第一节　再次程序化的遗传解码

一、翻译水平的移框

翻译水平的移框（translational frameshift）也称为翻译移框或核糖体移框，它是指核糖体在进行移位反应的时候，一次移动的核苷酸数目不是 3 个，而是 2 个或 4 个的现象。这种翻译水平的移框完全打破了原来的 ORF 的连续性，迫使翻译在一个全新的 ORF 内继续进行，直至遇到新的 ORF 内的终止密码子。通常将移动 2 个和 4 个核苷酸的移框分别称为 -1 移框和 +1 移框。

翻译中的移框经常出现在一些逆转录病毒（如 RSV 和 HIV-1 等）翻译 POL 蛋白（逆转录酶）的时候。RSV 的 *pol*、*gag* 以及 *env* 基因同处在一个多顺反子 mRNA 上，但由于在真核生物 mRNA 翻译起始的时候，起始密码子的识别依赖于 5′端的帽子结构，因此只有靠近 5′端的 *gag* 基因才能被翻译，而在后面的 *pol* 则无法正常地启动翻译。那么，如何让 *pol* 基因也能够翻译呢？有一种方法，就是让 *pol* 基因借用 *gag* 基因的起始密码子，乘 *gag* "翻译的东风"，接着翻译 *pol* 基因。然而如果借用，当翻译遇到 *gag* 的终止密码子，又怎么办？研究表明，核糖体在翻译到 *gag* 的终止密码子（UAG）前一个密码子处，进行了 -1 移框。终止密码子周围的核苷酸序列是 -ACA AAU UUA UAG GGA GGG CCA-，对应的氨基酸序列是 Thr-Asn-Leu。但经过 -1 移框以后，可读框变为 -ACA AAU UUAUA GGG AGG GCCA-，对应于的氨基酸序列是 Thr-Asn-Leu-Ile-Gly-Arg-Ala。于是，原来的 UAG 不能再作为终止密码子发挥作用了，核糖体继续翻译直到遇到 *pol* 基因的终止密码子。通过这种方式合成出来的先是 GAG 和 POL 的融合蛋白，然后通过特殊的蛋白酶的剪切，可释放出 POL 蛋白。以上发生移框的效率约为 5%，这意味着 95% 的翻译在遇到 *gag* 基因的终止密码子时将正常地终止，合成出来的产物为单独的 GAG 蛋白。这样一来，通过翻译移框，不但解决了 *pol* 基因不能正常翻译的问题，而且可以很好地满足病毒对于 GAG 蛋白和 POL 蛋白的不同需求。因为病毒对 GAG 蛋白的需求远远大于对 POL 蛋白的需求。

Quiz1 逆转录病毒的 POL 蛋白可以直接由核糖体启动翻译吗？为什么？

大肠杆菌 RF2 是通过 +1 移框得以翻译的例子（图 41-1）。已知 RF2 能识别 UAA 或 UGA 两种终止密码子，它本身的基因的第 26 个密码子为 UGA。如果细胞内的 RF2 的浓度较高，核糖体在翻译到这个密码子时就会终止。这时得到的产物并不是 RF2，而是一个无功能的 25 肽；如果细胞内的 RF2 的浓度较低，翻译就不能有效地终止。这时核糖体发生 +1 移框，翻译得以在一个新的可读框中继续进行，直到翻译出完整的 RF2。

Quiz2 若是将大肠杆菌RF2的第26个密码子突变成UAG，你预测将会对RF2的翻译移框有何影响？

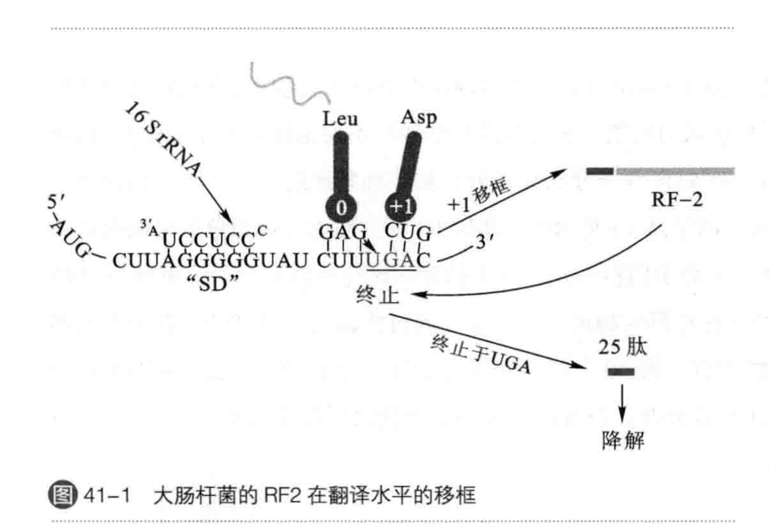


图41-1 大肠杆菌的RF2在翻译水平的移框

进一步的研究表明，发生移框的效率还是颇高的：30%的核糖体发生移框，得到全长的RF2。在移框位点，RF2-mRNA原来在核糖体P部位上的密码子（CUU）与肽酰-tRNA之间的反密码子（GAG）的配对被打破，然后在新的框架内，依赖与CUU重叠的密码子UUU和GAG之间的配对，使核糖体向前滑移了一个核苷酸的距离，从而使原来的UGA偏离了可读框，不再起终止作用。

实际上，在终止和移框之间始终存在着竞争，哪一种势力取胜，一是取决于RF2在细胞的浓度，二取决于移框位点的序列特征。RF2作为再解码的产物，其本身却促进核糖体终止于UGA或UAA，因此，当RF2量很多的时候，竞争倾向于终止而不是移框；反之，如果RF2的量供应不足，移框就在竞争中获胜，RF2的浓度恢复。移框位点本身的序列是有利于移框的：首先，UGA的后面是一个C，这是一个弱终止环境；其次，CUU本身是一个容易“打滑”的密码子。

更深入的研究发现，在移框位点的上游3个核苷酸的区域有一段刺激移框发生的序列。该序列与SD序列一样，能够与16S rRNA配对。于是，RF2-mRNA有两个SD序列：一个用于翻译的起始，另一个促进移框。移框位点和第二个SD之间的间隔序列长度非常关键，对于RF2来说，必须是3个碱基。第二个SD序列的作用可能是由于它与rRNA之间的配对，扭曲了翻译复合物，致使核糖体要么向前滑动发生正移框，要么向后滑动发生负移框。

鸟氨酸脱羧酶的抗酶（ornithine decarboxylase antizyme，ODA）是第一例在真核生物发现的发生+1移框的蛋白质（图41-2）。通过分析ODA的cDNA的碱基序列发现，其mRNA具有ORF1和ORF2的两个可读框，ORF2与ORF1重叠，并相差+1个核苷酸。但是，ORF1的碱基序列太短，不可能编码ODA，而ORF2的起始密码子无法识别故也不可能编码ODA。实验证明ODA的翻译是以ORF1的起始密码子开始，再通过移框进入ORF2继续翻译，最终得到目标产物，ORF2的移框需要UGA下游的假节结构。

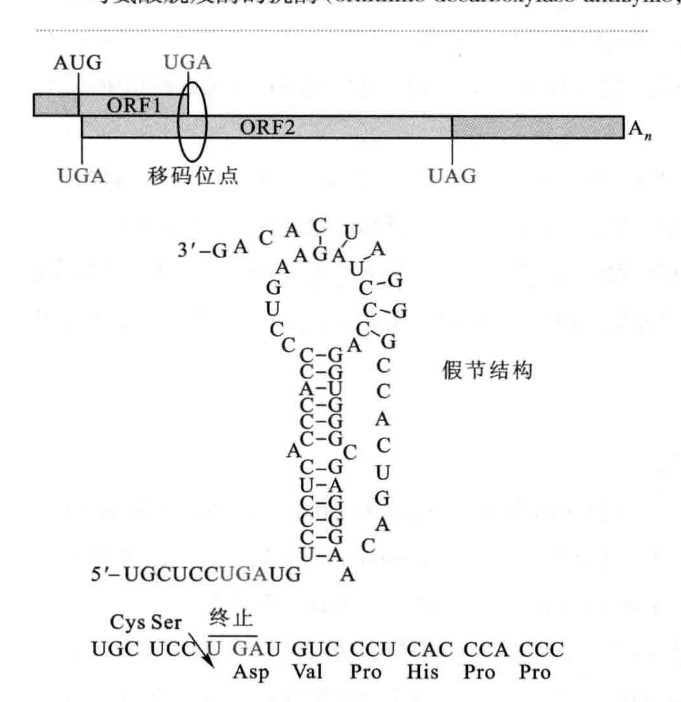


图41-2 鸟氨酸脱羧酶的抗酶在翻译水平的移框

二、通读

通读（read-through）是指翻译中将一个可读框内的终止密码子当做一个正常的氨基酸的密码子来阅读，而得到一个加长的多肽链的过程。存在这种现象的主要是一些病毒，如烟草花叶病毒（tobacco mosaic virus，TMV）和鼠类白血病病毒（MMLV）。MMLV是一种逆转录病毒，其POL蛋白也是与GAG蛋白一起作为融合蛋白被表达的，但并不是通过核糖体移框方式实现的，而是在翻译GAG的时候，经历了通读，将GAG的终止密码子UAG当做Gln的密码子。这里通读的效率也只有5%左右。

三、跳跃翻译

跳跃翻译(jumping translation 或 bypass translation)是指核糖体在翻译的过程中,直接跳过可读框中的一段核苷酸序列,再接着翻译下游序列的现象。这段漏读的序列并不决定任何氨基酸序列,被称为翻译水平的内含子(translational intron),以区别于分别在转录后加工和翻译后加工中除去的内含子和内含肽。对跳跃翻译了解得最深入的例子是 T4 噬菌体的基因 60 蛋白和大肠杆菌色氨酸阻遏蛋白(TrpR)。其中前者是美籍华裔学者黄惠敏和中国科学院上海生物化学研究所访问学者敖世洲于 1988 年共同发现的。基因 60 编码的是 DNA 拓扑异构酶的一个亚基。通过比较这个亚基的氨基酸序列和基因 60 的核苷酸序列发现,其可读框中有一段 50 个核苷酸组成的序列没有翻译,这段序列将前 46 个密码子(GGA)和后 114 个密码子(GGA)分开。没有证据证明基因转录后经历过剪接反应,唯一可能的解释是核糖体在翻译到第 46 个密码子后,发生了跳跃,跳过了这 50 个核苷酸序列组成的间隙(效率近 100%),继续翻译下游的 114 个密码子(图 41-3)。将基因 60 中长为 258 个核苷酸的片段与大肠杆菌的 β- 半乳糖苷酶(*lacZ*)基因构建成的融合质粒在大肠杆菌里表达,也出现跳跃现象。将基因 60 导入酵母细胞也发生跳跃。

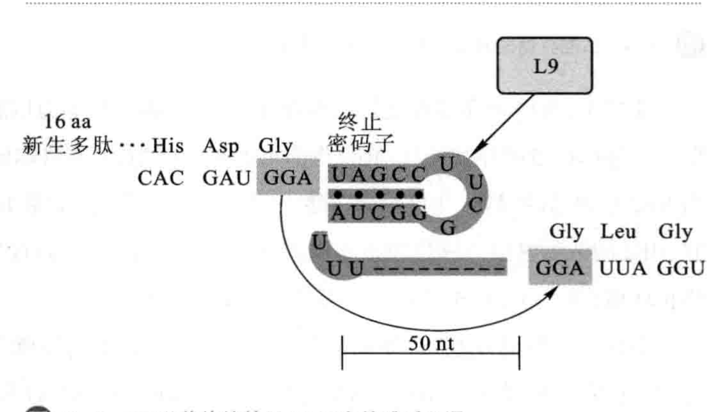


图 41-3 T4 噬菌体的基因 60 蛋白的跳跃翻译

TrpR 是跳跃翻译的又一个例子,该基因编码两种蛋白质,它们的相对分子质量分别是 12×10^3 和 10×10^3。这两种蛋白质 N 端氨基酸的序列完全是一样的,但 C 端的序列完全不同。经研究发现,10×10^3 是核糖体翻译 12×10^3 的可读框中发生了跳跃的结果,共跳跃 55 个核苷酸。由于 55 非 3 的整数倍,因此跳跃同时导致核糖体发生了移框,移框应属于 $55-3\times18=+1$ 移框。移框必然使得 10×10^3 蛋白质的 C 端氨基酸序列不同于 12×10^3 蛋白质。将含有 *trpR-lacZ* 融合基因的质粒在大肠杆菌中表达,也发生了跳跃。

至于核糖体为什么在特殊的位点发生跳跃,至今还没有一个令人信服的解释。在分析跳跃位点周围的碱基序列以及对这些序列进行点突变实验后,发现基因 60 核糖体的起跳点和着陆点处的密码子必须满足被同一种 tRNA 识别的条件。此外,被跳过的序列长度、其中的终止密码子和茎环结构、L9 对茎环结构稳定性的调节以及上游被翻译出来的 16 肽对跳跃均有影响,但是以上这些特征并不适用于 *TrpR* 基因。

四、含硒半胱氨酸和吡咯赖氨酸的参入

Sec 作为第 21 种蛋白质氨基酸,存在于一些特殊的蛋白质分子之中,如细菌的甲酸脱氢酶、哺乳动物的谷胱甘肽过氧化物酶、四碘甲状腺素 5′- 去碘酶(tetraikiodothyronine 5′-deiodinasem)和含有多个含硒半胱氨酸残基(7~10 个)的硒蛋白 P(selenoprotein-P),这些蛋白质统称为含硒蛋白。

Sec 的参入可视为一种特殊形式的通读,原因是含硒蛋白在翻译的时候,将 UGA 终止密码子作为 Sec 的密码子。那么,机体是如何将作为 Sec 的密码子——UGA,与作为正常终止密码子的 UGA 区分开的呢?对某些含硒蛋白基因的碱基序列所进行的分析以及点突变得到的结果表明:UGA 的通读与一些称为含硒半胱氨酸插入序列(selenocysteine insertion sequence,SECIS)的顺式元件有关。原核生物的 SECIS 的特征是在紧靠 Sec 密码子下游处有一个茎环结构,其中的茎由于非 Watson-Crick 碱基对(UGAC/UGAU)的存在而发生 90°的弯曲(图 41-4)。在大肠杆菌甲酸脱氢酶的 mRNA 分子上,位于 UGA 下游 40 bp 的区域也是 Sec 参入所必需的。

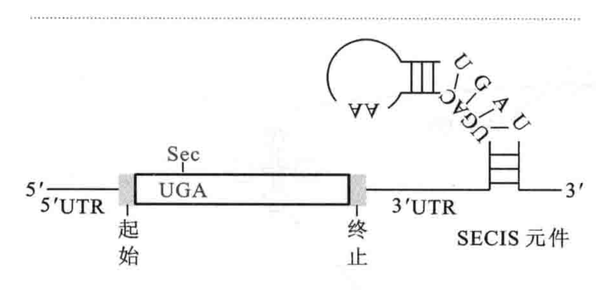


图 41-4 原核生物硒半胱氨酸密码子所处的环境

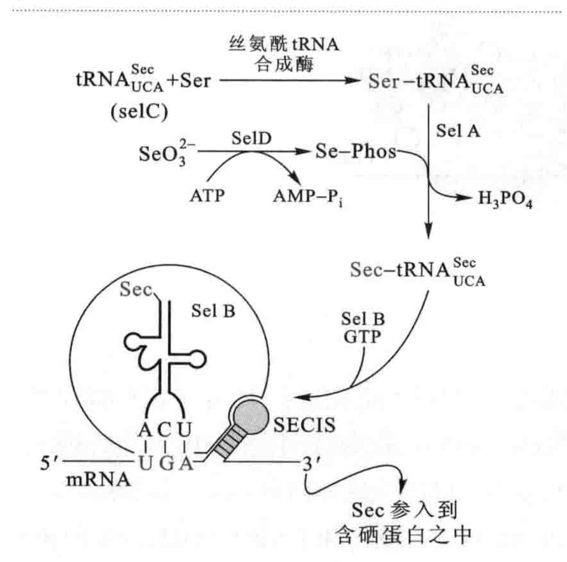


图 41-5 原核细胞含硒半胱氨酸的参入

真核生物的 SECLS 则存在于 3′端的非翻译区，在一级结构上无一致序列，但二级结构很相似。在四碘甲状腺素 5′－去碘酶 RNA 的 3′－UTR 中，有一段 250 bp 的区域（距离 UGA 约 1 kb）也是必需的。

大肠杆菌 Sec 的参入与 4 个基因的产物有关（图 41-5）：Sel A 将 Ser 转变成脱氢丙氨酸（dehydroalanine）；Sel B 为类 EF-Tu 延伸因子；Sel C 为 $tRNA^{Sec}$，其上的反密码子能识别 UGA；Sel D 激活 HSe^-。与其他 tRNA 相比，$tRNA^{Sec}$ 具有一些不同寻常的性质：含有 1 个 4 bp 环的 6 bp 茎环结构；一个加长的受体茎；一个大的附加臂。这些特征可阻止它与细胞中含量丰富的 EF-Tu 相互作用。在真核生物中，一种叫 SBP2 的蛋白质能特异性地与 SECIS 元件结合，在含硒半胱氨酸的参入中也发挥作用。

大肠杆菌 Sec-$tRNA^{Sec}$ 的形成通过下列反应：①Ser+$tRNA^{Sec}$→Ser-$tRNA^{Sec}$，该反应由丝氨酰 -tRNA 合成酶催化；②硒化物（selenide）+ ATP→磷酸硒化物 +AMP+Pi，该反应由磷酸硒化物合成酶（selenophosphate synthetase，为 SelD）催化；③磷酸硒化物 +Ser-$tRNA^{Sec}$→Sec-$tRNA^{Sec}$+H_2O，该反应由硒半胱氨酸合成酶（Sel A）催化。

Quiz3 严格地说，含硒半胱氨酸的名称并不确切，为什么？

当翻译进行到 Sec 的密码子 UGA 的时候，下游位置的 SECIS 与 Sec-$tRNA^{Sec}$、Sel B 和 GTP 形成四元复合物，Sec-$tRNA^{Sec}$ 进入 A 部位，肽酰转移酶催化 Sec 氨基的参入（图 41-6）。含硒半胱氨酸参入的效率并不总是 100%，例如在缺硒的条件下，哺乳动物的 1 型脱碘酶在翻译到 UGA 时会提前终止。

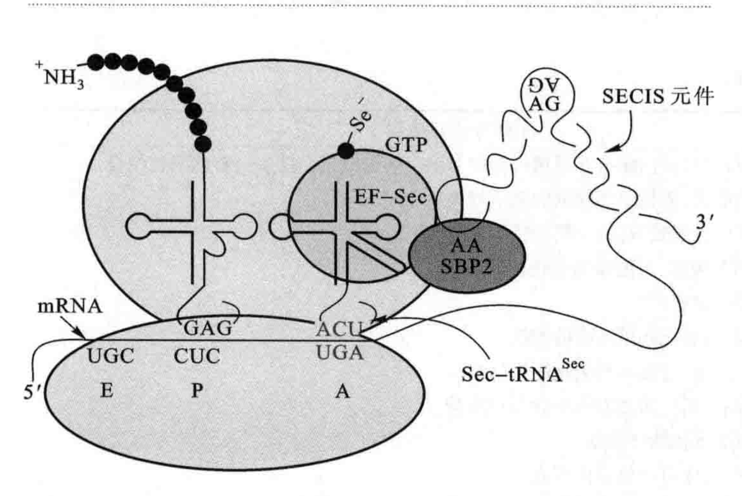


图 41-6 SECIS 元件在帮助 Sec-tRNASec 进入 A 部位中的作用

Pyl 是于 2002 年才发现的第 22 种蛋白质氨基酸，它的参入也是一种特殊的通读事件，发生通读的终止密码子为 UAG，由一种特殊的 $tRNA^{Pyl}$ 进行解码（*pylT* 基因编码）。这种 tRNA 带有的反密码子为 CUA。

Pyl-$tRNA^{Pyl}$ 的合成已被证明是由 *pylS* 基因编码一种第二类氨酰 -tRNA 合成酶催化，该酶直接催化游离的 Pyl 与 $tRNA^{Pyl}$ 形成 Pyl-$tRNA^{Pyl}$。

与 Sec 的参入一样，UAG 被通读成 Pyl 需要其下游的一段称为 Pyl 插入序列（pyrrolysine insertion sequencet，PYLIS）的元件，PYLIS 形成茎环结构，促进 Pyl 的参入（图 41-7）。

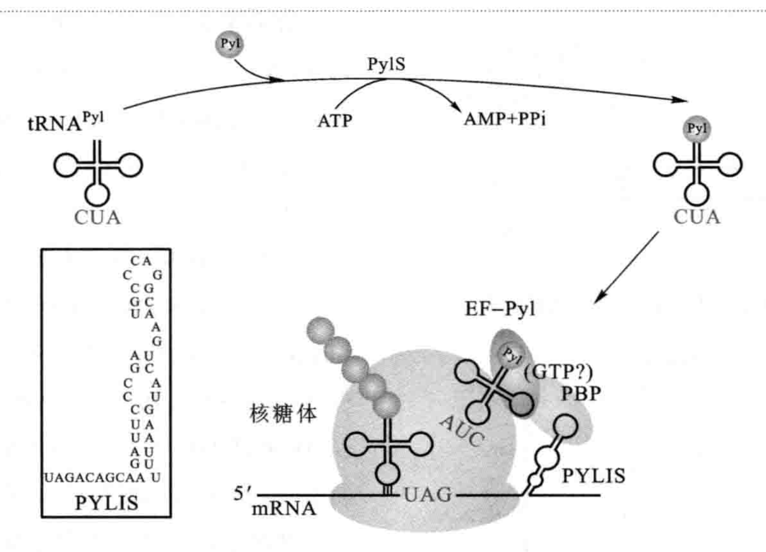


图 41-7 吡咯赖氨酸的参入

框 41-1 生化点滴——含硒蛋白知多少

硒是人体的必需微量元素，它对于健康来说是有益的，尤其对于免疫系统的功能、精子的游动和活性甲状腺素的形成都十分重要，此外，硒还能够降低得癌症和心血管疾病的风险，并能对 HIV 携带者发展为艾滋病有一定的抑制作用。但硒元素并不是越多越好，太多的硒可导致硒中毒(selenosis)。硒在机体中的作用是作为含硒半胱氨酸的成分，参入到含硒蛋白中，由含硒蛋白来完成。由于含硒半胱氨酸由终止密码子编码，因此缺硒可导致含硒蛋白的翻译提起终止。

人类食物中硒的主要来源是植物，但不可思议的是植物体内却没有含硒蛋白。1973 年，人们才发现了含硒蛋白。最早发现的三种是来自细菌的甘氨酸还原酶系统中的蛋白质 A 和甲酸脱氢酶，以及来自哺乳动物的谷胱甘肽过氧化物酶。

在哺乳动物体内，人们已发现近 25 种以上的含硒蛋白(表 41-1)，其中绝大多数在体内参与氧化还原反应。

► 表 41-1 哺乳动物体内的主要含硒蛋白

蛋白名称	序列特征与功能
含硒蛋白 15(Sep15)	位于内质网，与参与蛋白质折叠质量控制的 UDP- 葡糖糖蛋白糖基转移酶紧密结合
去碘酶	有 3 种不同的形式，催化甲状腺素的激活或失活
谷胱甘肽过氧化物酶	至少有 5 种不同的形式，有的呈组织特异性分布
Sel H	含有 CXXU 序列模体，可能参与氧化还原反应
Sel I	可能是一种膜内在蛋白
Sel K	膜蛋白，可能参与氧化还原反应的调控
Sel L	仅存在于海洋动物，可能参与氧化还原反应
Sel M	含有 CXXU 序列模体，可能参与氧化还原反应
Msr A	是一种甲硫氨酸硫化物还原酶
Sel N	存在于内质网，突变可导致肌肉疾病
Sel O	含有 CXXU 序列模体，可能参与氧化还原反应
Sel P	是血浆中的主要含硒蛋白，是已知唯一的一种含有多个 Sec 的蛋白质，例如人的 SelP 含有 10 个 Sec
Sel R	在细胞质或细胞核，保护细胞受氧化胁迫的破坏
Sel S	参与肝细胞的葡萄糖代谢和内质网蛋白的逆向归位
硒磷酸合成酶	催化硒化物和 ATP 形成单硒磷酸
Sel T	含有 CXXU 序列模体，可能参与氧化还原反应
硫氧还蛋白还原酶(TR)	有 TR1、TR2 和 TR3 三种形式，其中 TR1 存在于多种细胞的细胞质，参与氧化还原反应的调节，TR2 位于线粒体，TR3 位于睾丸的内质网，能还原 GSSG
Sel V	含有 CXXU 序列模体，可能参与氧化还原反应
Sel W	类似于 Sel V，只在肌肉组织中表达，可能具有抗氧化功能

第二节 翻译暂停

翻译的延伸并不一定以匀速进行，在 mRNA 编码区内可能存在一些“瓶颈”影响到延伸的速率，例如遇到稀有密码子，或者 mRNA 的某些区域的序列正好与 rRNA 序列同源。当核糖体遇到这些“瓶颈”的时候，翻译可能不得不暂停。也许，在编码区内发生的翻译暂停对于新生多肽链的正确折叠十分重要。在蛋白质共翻译的定向和分拣途径之中，SRP 与信号序列的结合也导致翻译的暂停，这样的暂停只有在 SRP/ 核糖体正确地停泊在内质网膜上以及 SRP 解离后才会解除。

翻译的暂停可让两个相邻的结构域独立地发生折叠。以酵母的一个与 Trp 合成有关的双功能酶为例，它由 *TRP3* 基因编码，其 N 端结构域具有邻氨基苯甲酸合成酶Ⅱ活性，而 C 端结构域具有吲哚甘油磷酸合成酶活性，其中的 C 端结构域含有 10 个相邻的稀有密码子。如果使用同义的富有密码子代替这 10 个稀有密码子，那对邻氨基苯甲酸合成酶Ⅱ的活性就无影响，但 C 端的吲哚甘油磷酸合成酶的比活性却降低了。这说明富有密码子的快速翻译，使得翻译出来的部分缺乏足够的时间进行最佳地折叠。

人体内有一种多药物抗性 1（multidrug resistance 1，MDR1）基因，编码一种叫 P- 糖蛋白（P-glycoprotein）的跨膜转运蛋白，能将多种药物运出细胞，从而影响药物疗效。在研究基因结构时，发现一种 *MDR1* 发生沉默突变的变体，与野生型相比，变体细胞在向胞外输出药物的效率上发生了显著变化。这种沉默突变虽然没有改变 P- 糖蛋白的一级结构，但却将它的一种氨基酸的富有密码子变成稀有密码子，从而导致翻译速度下降。这可影响到 P- 糖蛋白的折叠，使得突变体细胞内的 P- 糖蛋白的三维结构与野生型细胞有所不同，特别是在结合药物和抑制剂的位点上。这说明一级结构完全相同的蛋白质在不同细胞内可折叠成具有不同结构和功能的蛋白质。

此外，已有某些证据表明，N 端和 C 端的 α 螺旋或 β 折叠更喜欢使用富有密码子，而在 37 个大肠杆菌多结构域蛋白质的结构域之间的氨基酸残基更倾向使用稀有密码子。mRNA 的二级结构以及缺乏特定的 tRNA 也都可能导致翻译的暂停，而任何一种暂停有时可能会导致翻译的终止。

推荐网址：

e41-1 本章小结填填看

1. http://en.wikipedia.org/wiki/Translational_frameshift（维基百科有关翻译水平移框的内容）
2. http://bass.bio.uci.edu/~hudel/bs99a/lecture26/lecture7_6.html（美国加利福尼亚大学欧文分校一个生物学网站，有关含硒半胱氨酸参入蛋白质的内容）

参考文献：

1. Krzycki J A. Functional context，biosynthesis，and genetic encoding of pyrrolysine. *Current Opinion in Microbiology*，14：342–349，2011.
2. Kisselev L，et al. Translation termination in pyrrolysine-utilizing archaea. *FEBS Letters*，583：3 455–3 460，2009.
3. Ivanov I P，Atkins J F. Ribosomal frameshifting in decoding antizyme mRNAs from yeast and protists to humans：close to 300 cases reveal remarkable diversity despite underlying conservation. *Nucleic Acids Res*，35（6）：1 842–1 858，2007.
4. Vadim N，et al. Gladyshev，Pyrrolysine and Selenocysteine Use Dissimilar Decoding Strategies. *J Biol Chem*，280：20 740–20 751，2005.
5. Srinivasan G，et al. Pyrrolysine encoded by UAG in Archaea：charging of a UAG-decoding specialized tRNA. *Science*，296：1 459–1 462，2002.

第四十二章　原核生物的基因表达调控

任何一种生物的基因表达都受到严格的调控。无论是原核生物还是真核生物，基因组内的各基因在某一时刻并不需要都处于表达的状态，而即使表达的基因，其强度也不一样。例如，大肠杆菌基因组约有 4 000 个基因，但一般情况下只有 5%~10% 处于高表达状态，其他基因有的处于低表达状态，有的暂时不表达。实际上，所有生物体内的基因根据表达的状态可分为两组，一组是组成型基因（constitutive gene）或管家基因（house-keeping gene），这一组基因是维持细胞的基本活动所必需的，它们在所有的细胞内都自始至终处于开放状态，如参与糖酵解代谢的基因和核糖体蛋白基因；另一组是诱导型基因（inducible gene）或奢侈基因（luxury gene），它们仅仅在特定的细胞内、特定的生长或发育阶段或在特殊的条件下才会表达。不论管家基因还是奢侈基因，它们的表达都受到调控，只是调控的机制和幅度有所差别。

基因的表达模式可分为正调控（positive control）和负调控（negative control），其中原核生物主要使用负调控，而真核生物主要使用正调控。如果是在转录水平的调控，这两种调控模式一般都涉及特定的调节蛋白与 DNA 特定序列之间的相互作用。负调控是指在没有调节蛋白或者调节蛋白失活的情况下，基因正常表达；一旦有调节蛋白或者调节蛋白被激活，基因则不能表达。而正调控是指在没有调节蛋白或者调节蛋白失活的情况下，基因不表达或者低表达；一旦有调节蛋白或者调节蛋白被激活，基因才能表达或者大量表达。

原核生物通常生活在多变的环境之中，环境中的营养、温度、渗透压和 pH 等条件很容易发生变化，为了能够更好地生存和繁衍，需要随时改变自身基因的表达状况，通过选择性激活或抑制某些基因的活性，以调整体内执行相应功能的蛋白质或酶的种类和数量，从而改变细胞的代谢和活动。

原核生物基因的表达调控可以在几种不同的水平上进行，但最重要的控制环节是转录水平，尤其是在转录起始阶段。除此以外，基因表达调控也可以在 DNA 水平和翻译水平上进行。

本章将从几个不同的层次介绍原核生物基因表达的调控机制，会重点介绍操纵子、弱化子、反义 RNA 和核开关等重要概念。

第一节　在 DNA 水平上的调控

DNA 水平上的调控主要是通过基因的拷贝数、启动子的强弱和 DNA 重排来进行。

一、基因的拷贝数

一个基因的拷贝数直接影响到转录的效率，显然拷贝数越高，转录效率就越高。然而，细菌基因组 DNA 上的绝大多数基因为单拷贝，只有少数基因有多个拷贝，如 rRNA 的基因。rRNA 基因为多拷贝以及启动子是超强启动子，这两个性质可满足细胞对 rRNA 的大量需求。

二、启动子的强弱

这是调控管家基因表达的主要方式。管家基因需要每时每刻都表达，但不同的管家基因表达的效率并非一样，主要原因是不同的管家基因在启动子序列上不完全相同。在“DNA 转录”一章已强调过，一个基因的启动子序列与启动子的一致序列越接近，表达效率就越高，这样的启动子属于强启动

子；反之，表达效率就越低，属于弱启动子。对于某一种生物来说，其基因组内各个管家基因所具有的启动子是强是弱早已“命中注定”，在生物的进化过程中，它们在表达上的差异按照机体的需要已事先设定好。

Quiz1 你认为还有哪些机制可以让一个管家基因的表达“先天不足”？

三、DNA 重排

DNA 重排属于 DNA 重组的一种形式。通过重排，可以改变一个基因与其控制元件之间的关系，也可以缩短基因之间或基因片段之间的距离，从而实现对某些基因表达的调控。

蓝细菌（*Anabaena* sp.）PCC 7120 是一种丝状固氮菌。如果生活在含有复合氮（铵盐或硝酸盐）的培养基上，它们就会聚集在一起，形成一种仅由营养性细胞（vegetative cell）组成的长链状结构；如果没有复合氮，它们能将空气中的 N_2 固定成氨再加以利用。由于催化固氮反应的固氮酶遇到 O_2 就失活，因此，固氮反应只能发生在无氧的细胞中。

在没有复合氮的条件下，蓝细菌 PCC7120 能够分化出叫异胞体（heterocyst）的无氧细胞，以克服 O_2 对固氮反应的抑制。参与固氮的主要基因有 *nifH*、*nifD* 和 *nifK*，它们编码固氮酶的不同亚基。在营养性细胞的 DNA 上，这 3 个基因彼此之间相距甚远，特别是 *nifD* 和 *nifK* 相距约 11 kb。但在异胞体内，*nifD* 和 *nifK* 之间的间插序列已不见了。*nifD* 和 *nifK* 之间间插序列的消失，使 *nifD* 和 *nifK* 受控于同一个操纵子，从而能够等量表达（图 42–1）。

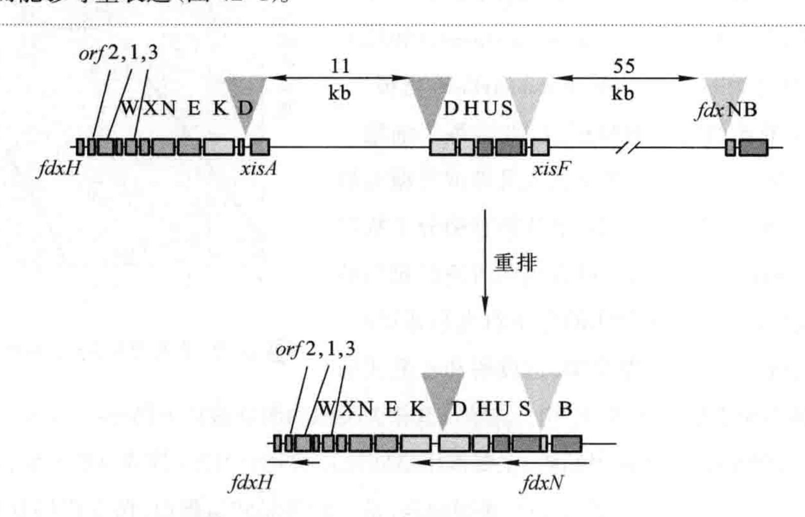


图 42–1 蓝细菌 PCC7120 异胞体内发生的基因重排对固氮基因表达的影响

那 *nifD* 和 *nifK* 之间的间插序列是如何被切除的呢？原来是异胞体在分化的过程中，经历了位点特异性重组，由 XisA 重组酶催化。XisA 重组酶利用了 *nifD* 和两侧由 11 bp 组成的直接重复序列进行重组。

第二节 在转录水平上的调控

原核生物的基因表达调控主要发生在转录水平上。绝大多数原核细胞的 mRNA 半衰期很短，因此，大多数多肽的翻译速率直接与其基因的转录活性相关联。一旦基因转录关闭，mRNA 就迅速被降解。

转录水平的调控既可以在转录的起始阶段，又可以在终止阶段进行，现分别加以介绍。

一、转录起始水平上的调控

操纵子（operon）是原核生物基因表达调控最重要的形式，原核生物的基因多数以操纵子的形式组成基因表达调控的单元。例如，大肠杆菌基因组共有 4 289 个基因，但绝大多数基因都被组织在约

578 个操纵子之中。

操纵子的概念和模型最早由法国科学家 Francois Jacob 和 Jacques Monod 于 1962 年提出。虽然真核生物一般没有操纵子结构，但操纵子所涉及的一些基本原理也适用于真核生物。

操纵子模型认为：一些功能相关的结构基因成簇存在，构成所谓的多顺反子(polycistron)，它们的表达作为一个整体受到控制元件(control element)的调节。控制元件由启动子、操纵基因(operator)和调节基因组成，其中操纵基因又称为操作子。调节基因编码调节蛋白，与操纵基因结合而调节结构基因的表达。如果调节基因编码的蛋白质与操纵基因的结合是阻遏基因的表达，这样的调控就为负调控，相应的调节蛋白就是阻遏蛋白(repressor)；如果调节基因编码的蛋白质与操纵基因的结合是激活基因的表达，这样的调控就为正调控，相应的调节蛋白就是激活蛋白(activator)。

(一) 大肠杆菌的乳糖操纵子

1. 乳糖操纵子的负调控

大肠杆菌的乳糖操纵子(*lac* operon)是第一个被阐明的操纵子。早在 20 世纪 50 年代，Jacob 和 Monod 就开始研究大肠杆菌的乳糖代谢，集中研究乳糖对乳糖代谢酶的诱导(induction)现象(图 42-2)：如果供大肠杆菌生长的培养基中没有乳糖，那细胞内参与乳糖分解代谢的 3 种酶，即β-半乳糖苷酶、乳糖透过酶(lactose permease)和巯基半乳糖苷转乙酰酶(thiogalactoside transacetylase)就很少，这时细胞内β-半乳糖苷酶的量很少，平均每个细胞只有 0.5~5 个。可是一旦在培养基中加入乳糖或乳糖类似物，在几分钟内，每个细胞中的β-半乳糖苷酶分子数量骤增，可高达 5 000 个，有时甚至可占细菌可溶性蛋白的 5%~10%。与此同时，其他两种酶的分子数也迅速提高。由此可见，新合成的β-半乳糖苷酶、透过酶和乙酰化酶是由底物乳糖或其类似物直接诱导产生，乳糖及其相关类似物因此被称为诱导物(inducer)。

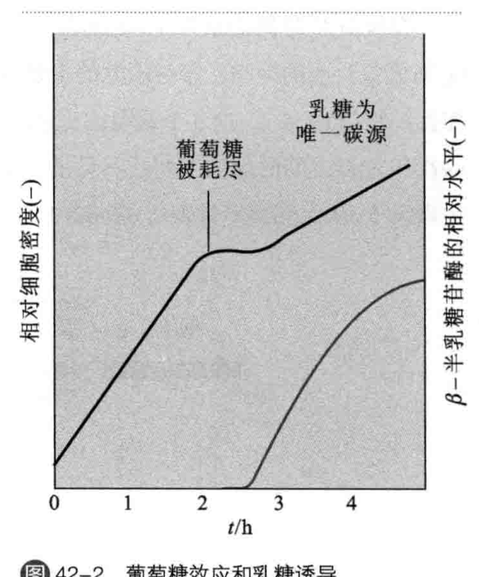


图 42-2 葡萄糖效应和乳糖诱导

β-半乳糖苷酶是由 *lacZ* 基因编码，主要作用是催化乳糖分子内β-糖苷键的水解，还催化少量乳糖变成别乳糖(allolactose)；透过酶由 *lacY* 基因编码，是一种跨膜转运蛋白，负责将培养基中的乳糖运输到胞内；乙酰化酶由 *lacA* 基因编码，可催化乙酰基从乙酰辅酶 A 转移到β-半乳糖苷上，但其功能至今仍不清楚。

为了搞清楚乳糖诱导现象的分子机制，Jacob 和 Monod 筛选出了一系列大肠杆菌乳糖代谢的突变体，并对各种突变体进行了详尽的遗传学研究，在取得翔实的实验数据的基础上，进行了严密的逻辑推理，最终提出了操纵子假说。

首先，他们注意到参与乳糖代谢的 3 个结构基因的调控是协调一致的，三种酶合成的时间和速度几乎相同，这就意味着它们可能有一个共同的控制元件或启动子；其次，遗传作图研究表明，三种酶的基因紧密连锁在一起，这为三种酶以多顺反子的形式存在提供了证据；随后的极性突变实验为多顺反子的存在提供了进一步的证据。他们的实验结果是：*lacZ* 的突变降低 *lacY* 和 *lacA* 的表达；*lacY* 的突变降低 *lacA* 的表达，但不影响 *lacZ* 的表达；*lacA* 的突变不影响 *lacZ* 和 *lacY* 的表达。

Jacob 和 Monod 还得到了其他一些很特别的突变体，在综合各种突变体的研究成果之后，终于提出了野生型大肠杆菌半乳糖操纵子的模型(图 42-3)。再经过多年不断的补充和改进，现代的乳糖操纵子模型已日臻完善，其主要内容是：①乳糖操纵子由调节基因、启动子、操纵基因和 3 个结构基因组成，其中调节基因、启动子和操纵基因构成控制元件，共同控制结构基因的表达。操纵基因位于启动子和结构基因之间，其核心结构是一段长为 21 bp 的回文序列；②调节基因(*lacI*)独立地表达，编码阻

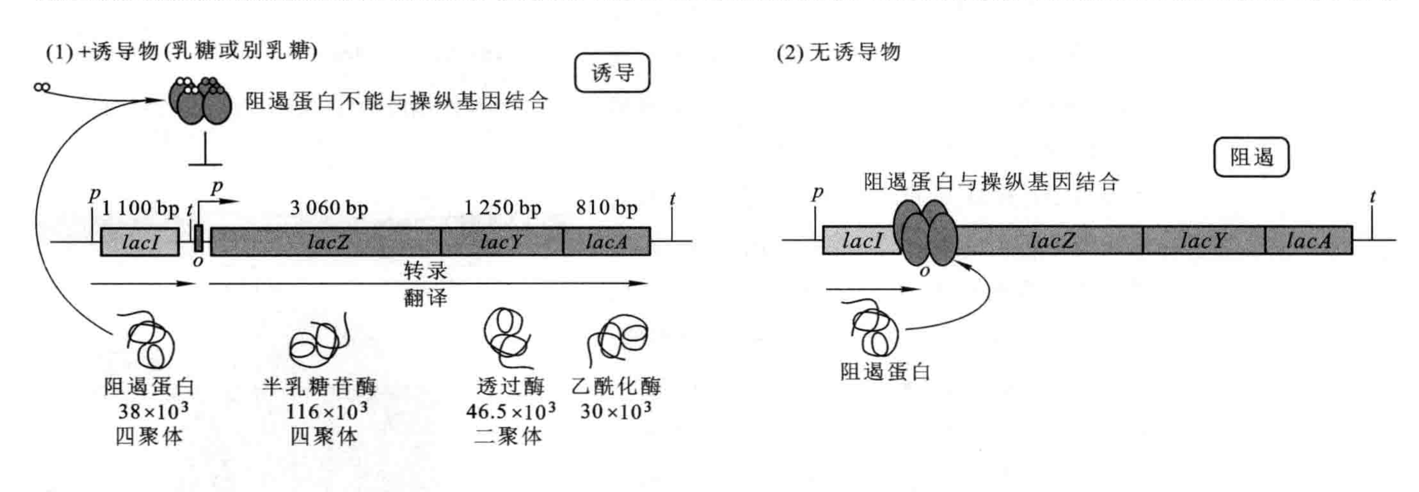


图 42-3 大肠杆菌半乳糖操纵子模型

遏蛋白，但由于受弱启动子的控制，阻遏蛋白在细胞内总是维持在较低的浓度；(3) 阻遏蛋白为四聚体蛋白，在无乳糖的情况下，与操纵基因（*lacO*）结合，阻断 RNA 聚合酶启动结构基因的转录，但这种结合并不完全，因此，会有微量的 β- 半乳糖苷酶、乳糖透过酶和转乙酰酶的合成；④一旦高浓度的乳糖进入细胞，在细胞内残留的 β- 半乳糖苷酶催化下，一部分乳糖发生异构化，变成别乳糖。而别乳糖作为别构效应物与阻遏蛋白结合，改变阻遏蛋白的构象，使其不能再与操纵基因结合，操纵子就被打开；⑤在阻遏蛋白与操纵基因解离以后，RNA 聚合酶与启动子结合，启动 3 个结构基因的转录，产生 *lacZ*、*lacY* 和 *lacA* 的共转录物，但翻译却是独立地进行，从而产生 3 种不同的酶；⑥由于阻遏蛋白与操纵基因的结合阻断了结构基因的表达，因此，乳糖操纵子受到它的负调控；⑦发生在控制元件内的突变可影响到结构基因的表达。

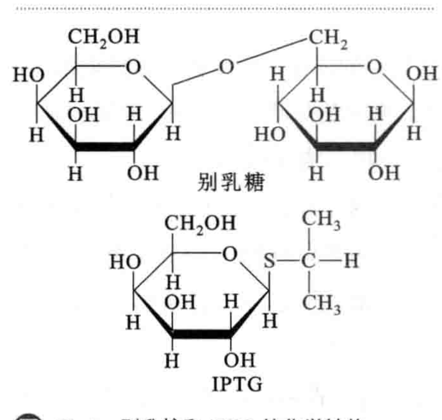


图 42-4 别乳糖和 IPTG 的化学结构

异丙基硫代 -β-D- 半乳糖苷（isopropyl-D-thiogalactopyranoside，IPTG），是一种人工合成的乳糖类似物（图 42-4），能够迅速和持续地刺激乳糖操纵子结构基因的表达，这是因为它本身并不能被 β- 半乳糖苷酶降解，因此被称为安慰诱导物（gratuitous inducer）。

X 射线衍射数据表明，构成乳糖操纵子的阻遏蛋白的每一个亚基上共有 3 个结构域：N 端结构域（1~62）、核心结构域（63~340）和 C 端结构域（341~357）。①N 端结构域还可以分为 2 个功能区，一个与 DNA 上特殊的碱基序列结合（1~45），另一个为铰链区（46~62）。与 DNA 结合的区域含有一种典型的 DNA 结合模体——螺旋 - 转角 - 螺旋，它主要在双螺旋大沟与 DNA 结合。铰链区将阻遏蛋白的 DNA 结合区域与核心结构域联系起来，使得两个结构域可以独立地移动。此外，铰链区也参与阻遏蛋白与操纵子的结合，其内部的一段无规则卷曲（50~58）在与 DNA 结合以后转变成 α 螺旋，这有助于 DNA 结合区域采取更好的取向，从而稳定阻遏蛋白与操纵基因的结合。②核心结构域为诱导物结合区域，由两个结构相似的亚结构域组成。每一个亚结构域含有一个 6 股 β 折叠和环绕在外的 4 段 α 螺旋。③C 端结构域与四聚体的形成有关。

乳糖操纵子实际上含有 3 个阻遏蛋白结合位点——O_1、O_2 和 O_3（图 42-5）。这 3 个位点是实现最大限度的阻遏所必需的。O_1 就是最早发现的操纵基因序列，它与启动子部分重叠；O_3 位于 *lacI* 的内部，在 O_1 上游 93 bp 的区域；O_2 位于 *lacZ* 的内部，在 O_1 下游 401 bp 的区域。这三个位点都具有接近二重对称的结构，其中 O_1 要比 O_2、O_3 对称，因此与阻遏蛋白结合最牢。在阻遏蛋白 -DNA 复合物之中，每一个阻遏蛋白四聚体与 2 个独立的结合位点结合，即每一个二聚体与 1 个结合位点结合。

当 3 个位点都存在的时候，阻遏蛋白抑制转录起始的幅度可达近 1 000 倍；如果缺乏 O_2 或 O_3，阻遏效果就只有约 500 倍；如果 O_2 和 O_3 都缺乏，阻遏效果就只有 20 倍。若阻遏蛋白结合到 O_1 和 O_3 上，O_1 和 O_3 之间的 DNA 就弯曲成环，从而进一步降低 RNA 聚合酶与启动子的结合。若阻遏蛋白结合到 O_1 和 O_2 上，O_1 和 O_2 之间的 DNA 就弯曲成环，这可直接干扰 mRNA 的合成或延伸。

一旦诱导物在核心结构域与阻遏蛋白结合，阻遏蛋白的构象即发生变化，并迅速通过铰链区传播到 DNA 结合区域，而导致铰链区内的 α 螺旋结构的破坏，进而降低阻遏蛋白与操纵基因之间的亲和力，致使 DNA 被释放出来。

2. 乳糖操纵子的正调控

乳糖操纵子除受到阻遏蛋白的负调控以外，还受到分解物激活蛋白（catobolite activator protein，CAP）的正调控。

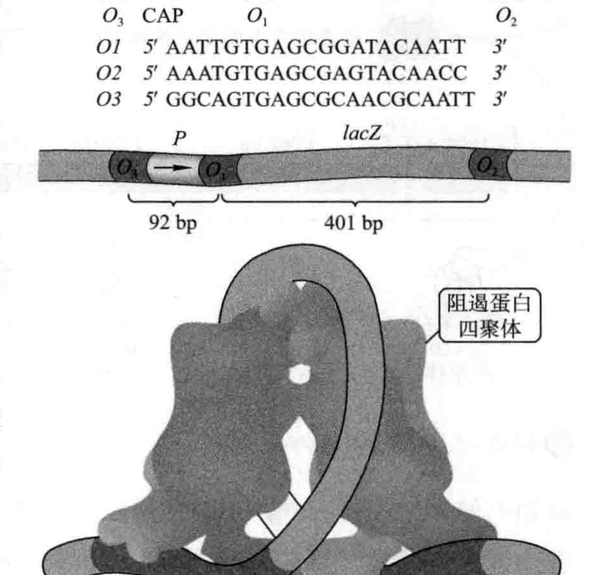


图 42-5 乳糖操纵子阻遏蛋白与三个结合位点的结合以及 DNA 环结构的形成（Watson 等，2007）

正调控是在对大肠杆菌的另外一种代谢现象，即葡萄糖效应（glucose effect）的研究中发现的。人们很早就知道，葡萄糖的存在能够阻止大肠杆菌对其他糖类的利用，这种现象被称为葡萄糖效应。例如，当大肠杆菌在同时含有葡萄糖和乳糖的培养基中生长的时候，大肠杆菌并不能利用乳糖，只有当葡萄糖被耗尽以后，细胞才能够代谢乳糖。那么，葡萄糖是如何抑制细胞利用乳糖的呢？

1965 年，B Magasonik 等人发现，在大肠杆菌中也有 cAMP，而且它的浓度与葡萄糖浓度呈负相关。cAMP 浓度的变化与腺苷酸环化酶（AC）的活性直接相关联，即高浓度的葡萄糖因抑制 AC 的活性而导致 cAMP 浓度的下降。那么，是不是细胞内 cAMP 浓度的下降致使乳糖不能被利用的呢？

为了弄清 cAMP 浓度的变化与乳糖利用之间的关系，有人将易于通过大肠杆菌细胞膜的 cAMP 类似物——双丁酰 -cAMP 加入到含有葡萄糖和乳糖的培养基中，结果发现乳糖操纵子受到诱导，乳糖能够被利用了。这就说明 cAMP 浓度的升高的确是细胞能够利用乳糖的前提。然而，cAMP 浓度的升高又是如何打开乳糖操纵子的呢？

当人们得到两种很特殊的大肠杆菌突变体以后，上面的问题才有了答案。这两种突变体都只能利用葡萄糖，不能利用其他碳水化合物，其中一种突变体的 *AC* 基因有缺陷，因此在任何情况下，都不能合成 cAMP；另一种突变体缺乏一种能与 cAMP 结合的蛋白，即 cAMP 受体蛋白（cAMP receptor protein，CRP）。CRP 就是分解物激活蛋白 CAP，这种突变体在加入外源的 cAMP 以后也不能利用乳糖。这两类突变体的存在不但进一步确定了 cAMP 与乳糖代谢之间的关系，而且说明了 cAMP 是通过 CAP 起作用的。

CAP 由 2 个相同的亚基组成，每 1 个亚基有 2 个结构域，1 个在 N 端，含有 cAMP 结合位点，另 1 个在 C 端，含有螺旋 - 转角 - 螺旋，负责与 DNA 结合。CAP 必须与 cAMP 结合以后才有活性。当 cAMP 与 CAP 结合以后，CAP 的构象即发生变化，其 C 端的螺旋 - 转角 - 螺旋采取合适的取向，从而能够识别并结合 DNA 上的特异性位点。

CAP-cAMP 与 DNA 结合的特异性位点由 26 bp 组成，位于乳糖操纵子的启动子紧靠 -35 区的上游，其一致序列是一段不完善的回文序列——TGTGA-N_6-TCACA，该位点称为 CAP 位点。CAP-cAMP 与 CAP 位点的结合可导致周围的 DNA 产生小的弯曲，CAP 因此能够与 RNA 聚合酶全酶在 α 亚基

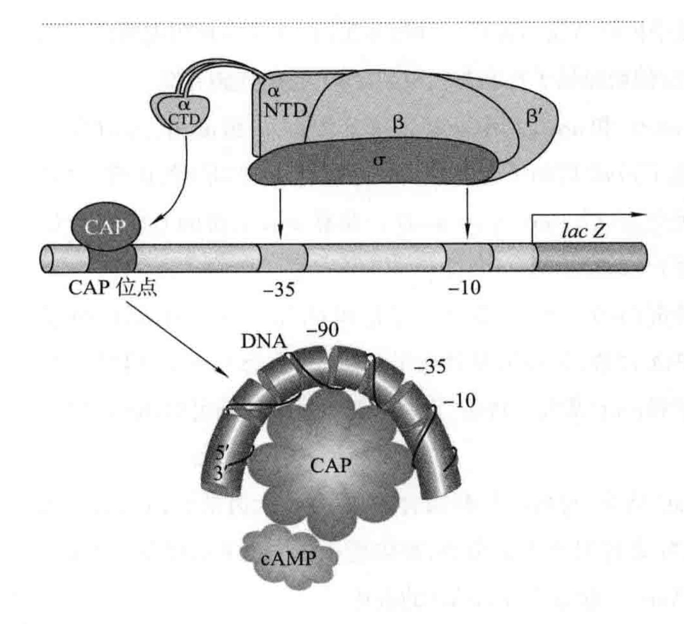


图 42-6　CAP-cAMP 与 RNA 聚合酶以及乳糖操纵子的相互作用（Watson 等，2007）

的 C 端结构域（α-CTD）相互作用，从而有利于 RNA 聚合酶与启动子的结合，以及 DNA 双螺旋的局部解链，最终激活了下游基因的转录（图 42-6）。

CAP 位点不仅存在于乳糖操纵子的附近，还存在于其他一些与碳源分解代谢有关的操纵子的周围，如半乳糖操纵子和阿拉伯糖操纵子，这也说明 CAP-cAMP 能激活多个操纵子的活性。

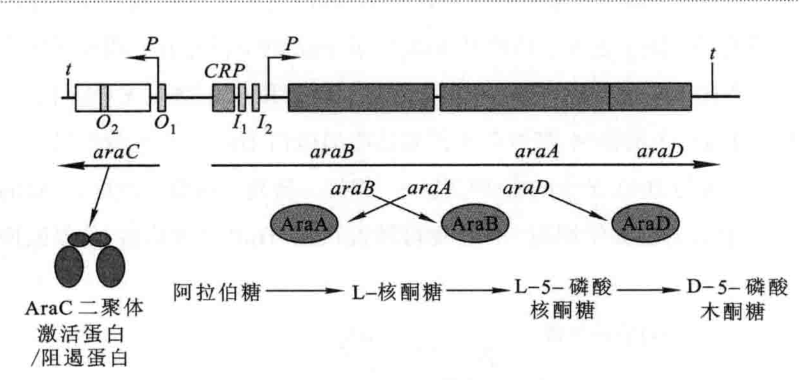
乳糖操纵子之所以受到双重调控，有 2 个原因：一是使细胞能够优先利用葡萄糖，而优先利用葡萄糖对细胞来说是有益的，因为参与葡萄糖分解的即编码糖酵解的各个酶的基因均是管家基因，这样一来葡萄糖可以迅速地被分解，为细胞提供能量；二是 *lac* 启动子天生是一个弱启动子，CAP-cAMP 的激活就弥补了其启动子活性的“先天不足”。

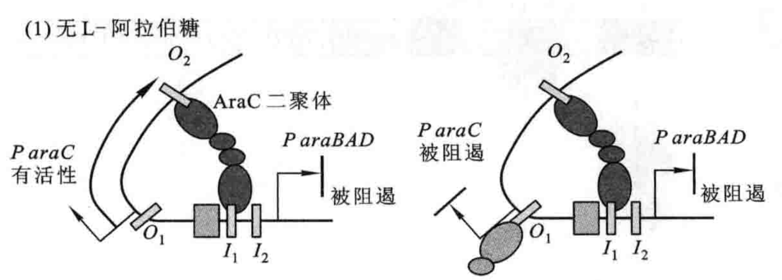
（二）大肠杆菌的阿拉伯糖操纵子

乳糖操纵子模型的巨大成功曾使 Monod 一度认为，大肠杆菌内所有基因表达的调控都是负调控。然而，就在他与 Jacob 一起建立乳糖操纵子模型的时候，Ellis Englesberg 却在研究大肠杆菌对阿拉伯糖的代谢。Englrsberg 在进行了一系列遗传学分析以后，认为阿拉伯糖操纵子受到正调控。他的正调控假说在相当长的一段时间以后才被接受。

Quiz2 如果将乳糖操纵子的启动子 -10 区序列突变成 TATAAT，你认为还会有葡萄糖效应吗？为什么？

阿拉伯糖操纵子编码 3 个与阿拉伯糖代谢有关的酶（图 42-7）：由 *araA* 基因编码的阿拉伯糖异构酶，由 *araB* 基因编码的核酮糖激酶，由 *araD* 基因编码的 5- 磷酸核酮糖差向异构酶。这 3 个结构基因按照 *araB*、*araA* 和 *araD* 的顺序排列，简称为 *araBAD*，共同受 *araC* 基因的产物 AraC 蛋白和 CAP-cAMP 控制。





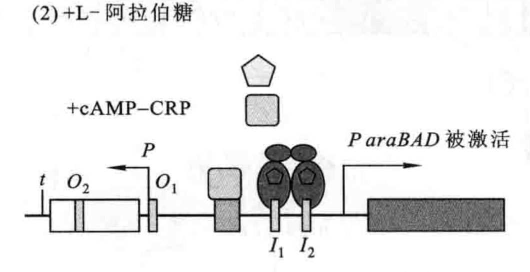


图 42-7　大肠杆菌阿拉伯糖操纵子模型及其作用机制（Weaver，2011）

与乳糖操纵子不同的是，阿拉伯糖操纵子的调节蛋白既是一种激活蛋白，又是一种阻遏蛋白。那AraC是如何做到这一点的呢？仔细观察阿拉伯糖操纵子在调节区域的结构，就能找到答案。

AraC有3个不同的结合位点：$araO_2$、$araO_1$和$araI$，其中$araI$又可分为$araI_1$和$araI_2$。$araI$位于CAP位点和*araBAD*的启动子之间；$araO_1$位于*araC*启动子的上游；$araO_2$位于*araC*基因的内部，远离$araO_1$。AraC与3个结合位点的亲和力依次是$araI >> araO_2 >> araO_1$。紧靠$araI$上游的CAP结合位点用来激活*araBAD*启动子，但对*araC*启动子无效。

在无阿拉伯糖的时候，AraC作为阻遏蛋白与I_1和O_2结合。与I_1和O_2结合的AraC蛋白分子相互作用形成二聚体，导致它们之间的DNA片段形成环，从而阻止RNA聚合酶与*araC*启动子和*araBAD*启动子的结合，抑制阿拉伯糖操纵子和*araC*基因的转录。此外，过量的AraC还可以和O_1结合，而进一步抑制自身的合成。

Quiz3 你认为使用随机突变的方法获得组成型的乳糖操纵子，与获得组成型的阿拉伯糖操纵子相比，哪一个更容易？为什么？

在有阿拉伯糖的时候，阿拉伯糖与AraC结合，与阿拉伯糖结合的AraC作为激活蛋白与I_1和I_2结合，原来在I_1和O_2之间的DNA环消失，如果这时也无葡萄糖，那细胞内的cAMP浓度就会升高，cAMP-CAP结合在I_1附近的CAP位点，与AraC一起激活AraBAD的表达。

（三）大肠杆菌的色氨酸操纵子

乳糖操纵子和阿拉伯糖操纵子都属于诱导型操纵子(inducible operon)，这是因为乳糖和阿拉伯糖都是作为诱导物来诱导各自操纵子发生转录的。下面要介绍的色氨酸操纵子是一种阻遏型的操纵子(repressible operon)，它控制5种参与色氨酸合成的酶的基因的表达。之所以说它是阻遏型操纵子，是因为色氨酸作为辅阻遏物(co-repressor)，与阻遏蛋白结合而阻止色氨酸操纵子的表达。

除了色氨酸操纵子是阻遏型以外，其他与合成代谢有关的操纵子也属于阻遏型操纵子，如组氨酸操纵子。一般说来，控制分解代谢的操纵子为诱导型，控制合成代谢的操纵子属于阻遏型。操纵子如此的分工使得细胞能够对环境或胞内的代谢变化迅速做出反应。

以色氨酸操纵子为例，如果培养基中含有Trp，那大肠杆菌就只需将它运输到细胞内直接利用，而不必浪费能量自己去合成，因此，这时候色氨酸操纵子应该被关闭；相反，如果培养基中没有Trp，大肠杆菌就必须"自力更生"，自己去合成，以满足蛋白质合成对它的需求，这时候色氨酸操纵子应该被打开。

色氨酸操纵子的结构如图42-8所示，其结构基因包括*trpE*、*trpD*、*trpC*、*trpB*和*trpA*，这些结构基因编码的酶都与色氨酸的合成代谢有关。除了这5个结构基因以外，在*trpE*的上游还有一段前导序列(*trpL*)，编码一段寡肽，其内部含有弱化子序列(见后)；操纵基因序列在启动子和*trpL*之间，无CAP位点；调节基因*trpR*远离操纵子，受自身启动子的控制，持续低水平表达阻遏蛋白TrpR。与乳糖操纵子的阻遏蛋白和CAP一样，TrpR也含有可与DNA结合的模体结构——螺旋-转角-螺旋。然而，单独的TrpR并不能与色氨酸操纵基因结合，因此，如果细胞内的色氨酸浓度很低，TrpR就无活性，色氨酸操

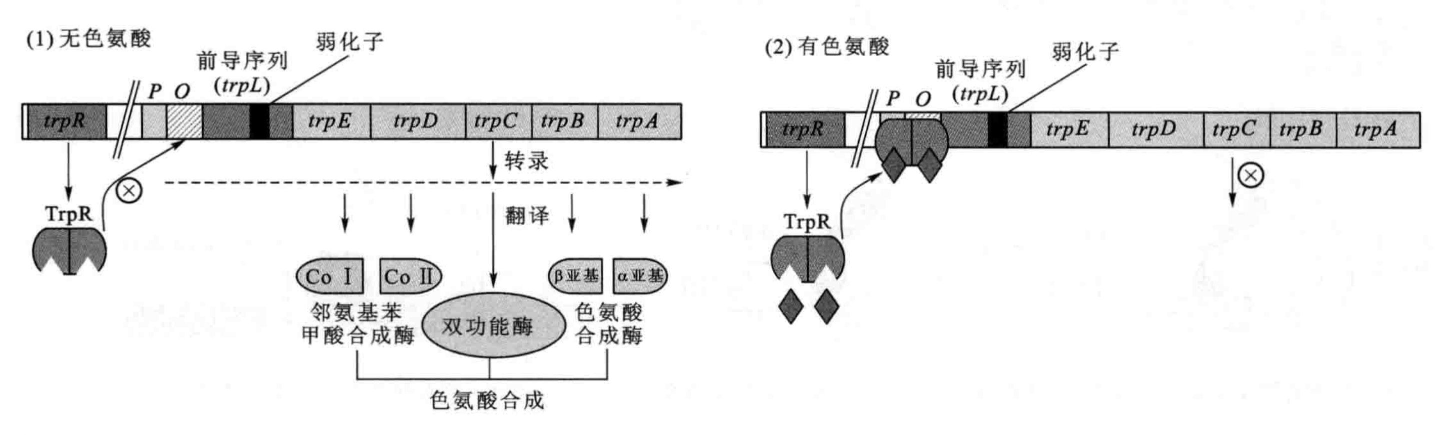


图42-8 大肠杆菌色氨酸操纵子模型

纵子会处于开放的状态,参与色氨酸合成的酶就会表达,为细胞合成必需的 Trp。当细胞内的 Trp 积累到一定水平以后,色氨酸作为辅阻遏物与 TrpR 结合,诱导其构象变化使之激活,随后 TrpR–Trp 就与操纵基因结合,而阻遏结构基因的转录。

参与色氨酸合成的 5 个基因除受到操纵子调控以外,还受到一种更为精细的调节方式即弱化子的调控,详情参看"弱化子"。

(四) 可变 σ 因子的切换使用

细菌负责识别启动子序列的是 RNA 聚合酶全酶中的 σ 因子,但细胞内的 σ 因子并非只有一种,许多细菌除了具有一种主要的 σ 因子负责识别绝大多数基因的启动子以外,还具有一些特别的 σ 因子,负责识别一些在特殊生长条件下(如热休克、氮饥饿、孢子形成和细菌生长稳定期适应)才需要表达的基因(参看第三十六章"DNA 转录"表 36–2)。

许多噬菌体合成自己的 σ 因子,以改变宿主细胞 RNA 聚合酶对启动子的特异性,从而选择性地表达自己的基因,而枯草杆菌的 SPO1 噬菌体则是通过 σ 因子的级联表达,来控制其不同感染阶段的基因表达(图 42–9)。

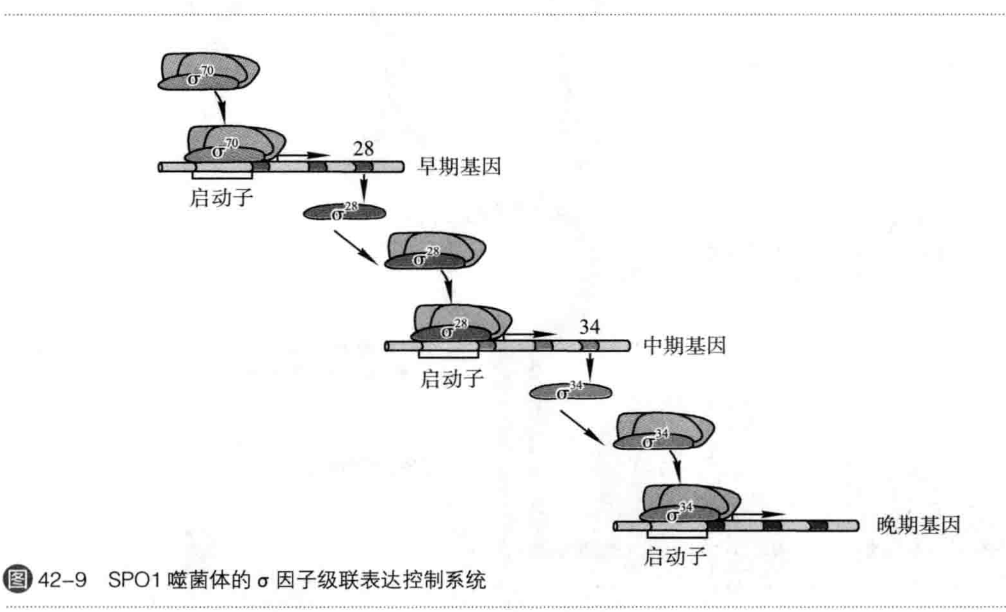


图 42–9 SPO1 噬菌体的 σ 因子级联表达控制系统

二、转录终止水平上的调控——终止与抗终止

与转录的起始相比,转录的终止具有较大的灵活性。通过改变基因转录的终止,不仅可以调节一个终止子下游的基因是否表达,还可以影响到一种转录物 3′UTR 的结构,从而进一步在转录后水平对基因表达实行调控。

有三种调节转录终止的手段:一是弱化(attenuation);二是核开关;三是抗终止。

(一) 弱化

此手段是建立在原核生物的转录与翻译偶联的基础上,通过调节特殊的终止子活性来进行的。

1. 大肠杆菌色氨酸操纵子内的弱化子

如果色氨酸操纵子是控制 Trp 合成有关酶基因表达的唯一手段,就不难设想,TrpR 的缺失将导致色氨酸操纵子对培养基中加入的 Trp 不再敏感,因为细胞内已没有感应色氨酸浓度变化的装置。然而,事实却大大出乎人们的意料,加入的 Trp 仍然能够降低 *trpEDCBA* 基因的表达,这就表明了细胞内一定存在其他的机制调节 *trpEDCBA* 基因的表达。

进一步的研究显示,第二种调节机制就是参与弱化的弱化子(attenuator),它与 tRNA(特别是色氨酰 –$tRNA^{Trp}$)和 *trpL* 基因有关。

弱化子是一种更为精细的调节基因表达的模式,它建立在翻译和转录之间偶联的基础上,因此为

原核细胞所特有。弱化子一般存在于参与生物合成的操纵子之中，与操纵基因一起，共同调节参与生物合成酶的基因的表达。

如图 42-10 所示，大肠杆菌的色氨酸的弱化子序列位于其操纵子的 *trpL* 之中。*trpL* 位于操纵基因和 *trpE* 之间，其内部含有一个小的 ORF——编码一个由 14 个氨基酸残基组成的前导肽。这个小 ORF 内部含有 2 个连续的色氨酸密码子，作为细胞又一种检测内部色氨酸水平高低的装置。色氨酸若供应充足，那就很容易被运载到 $tRNA^{Trp}$ 分子上，形成色氨酰 $-tRNA^{Trp}$，前导肽的翻译就不成问题；但色氨酸若供应不足，色氨酰 $-tRNA^{Trp}$ 就难以形成，这时前导肽的翻译就会停留在色氨酸密码子处，不能继续翻译。然而，“前导肽能否正常翻译”是如何转化成“基因能否继续转录”的结果呢？

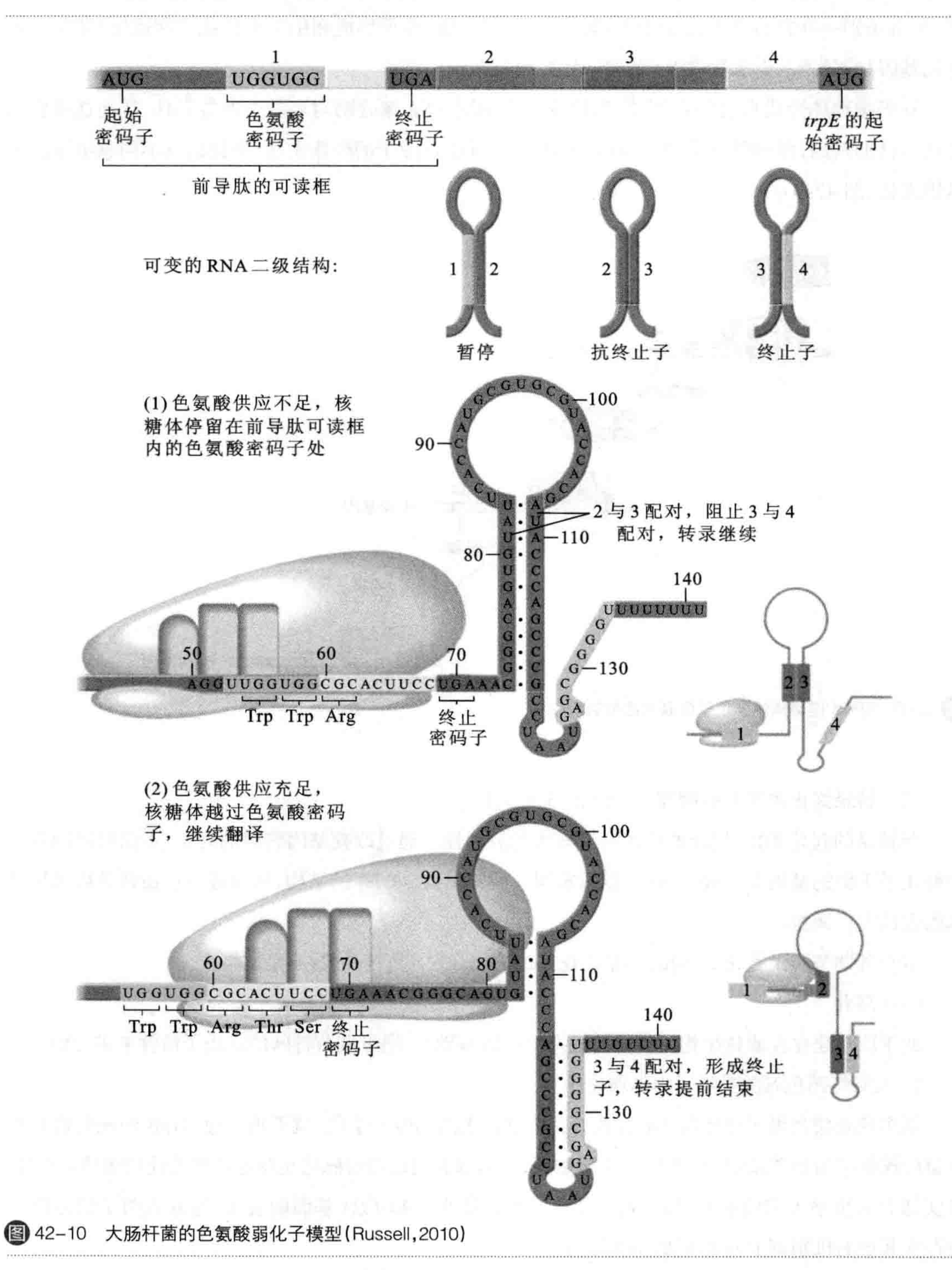


图 42-10　大肠杆菌的色氨酸弱化子模型（Russell，2010）

在获得 *trpL* 全序列以后，人们对整个区域可能形成的二级结构进行了预测。结果表明，该区域含有 4 段特别重要的核苷酸序列，按照 5′→3′的方向依次编号为 1、2、3 和 4，这 4 段 RNA 有 2 种配对方式，一种是 1 与 2、3 与 4 之间配对，另一种是 2 与 3 之间配对，不同的配对就形成不同的茎环结构。如果

是 3 与 4 配对，就形成转录的终止子结构；如果是 2 与 3 配对，形成的茎环结构就不是终止子。因此，当将前导链 mRNA 可能形成的二级结构与细胞内 Trp 水平的高低联系在一起的时候，就不难理解弱化子作用的妙处了。

由于原核生物的转录和翻译是偶联的，因此 *trpL* 一旦开始转录，就会有核糖体结合上来，并翻译其中的 ORF。如果细胞内有充足的色氨酸，翻译就一直会持续下去，直至遇到终止密码子。前导肽的顺利翻译致使 2 和 3 之间不能配对，但 3 和 4 之间可以配对而形成终止子结构（小的茎环结构）。终止子结构一旦出现，*trpEDCBA* 基因的转录提前结束。相反，如果细胞内的色氨酸供应不足，核糖体就会暂停在 ORF 内色氨酸密码子之处，等待色氨酰 -tRNA^{Trp} 的进入。前导肽的翻译不畅致使 2 和 3 之间配对形成大的茎环结构（非终止子），但 3 和 4 之间却不能配对，由于没有终止子，*trpEDCBA* 基因可继续转录下去。

弱化子的存在，使得“逃过”阻遏这一关的非正常转录得到及时“刹车”，这对操纵子的阻遏效应是一个很好的补充。而对那些无阻遏蛋白的操纵子来说，弱化子成为控制结构基因表达的唯一手段。

Quiz4 如果将大肠杆菌色氨酸操纵子前导肽内编码色氨酸的密码子分别增减一个，那么将会对弱化子的作用带来什么影响？

据估计，色氨酸操纵子的阻遏可实现 80 倍的调控，而弱化子可将调控再提高 6~8 倍，综合起来调控的幅度可达 500 倍左右。

与色氨酸操纵子一样，其他氨基酸操纵子在 5′端都具有类似的编码前导肽的小 ORF，而在每一个小 ORF 的内部，总会有几个连续的编码相应氨基酸的密码子。

（二）核开关

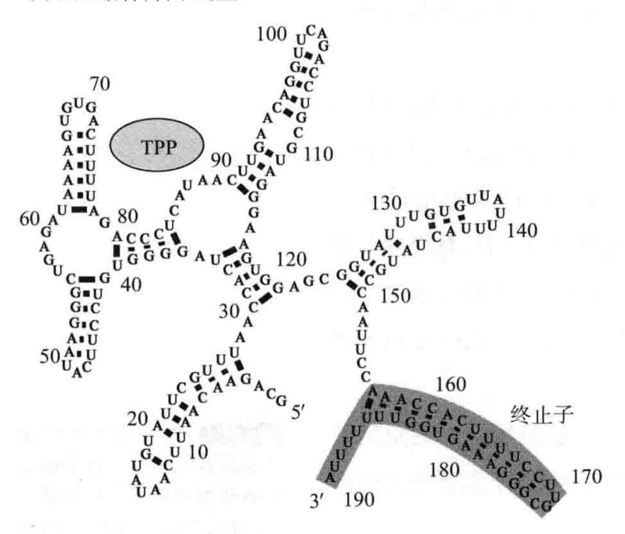
每一个细胞必须对其内外环境的变化做出反应，从而改变一系列基因的表达。然而，在细胞对内外环境变化做出任何反应之前，都需要存在一种机制来检测发生的变化。在操纵子一节，我们已经看到阻遏蛋白或激活蛋白具有这样的功能，如乳糖操纵子的阻遏蛋白能够监测到细胞内乳糖水平的变化，CAP 能够监测到细胞内 cAMP 浓度的变化。自 2000 年以来，在许多细菌和某些真核生物（如真菌和拟南芥等）中，已发现一些特殊的双功能 mRNA 具有类似的功能，这些特别的 mRNA 在非编码区含有特定代谢物或者金属离子（如 Mg^{2+}）的特异性结合位点，这些特异性的结合位点有时被称为适配体或适体（aptamer）。在这里，适配体或适体充当一种基因表达的开关，代谢物与其结合可改变 mRNA 的构象，结果要么提高转录的终止效率，要么降低翻译的效率，要么影响到 mRNA 后加工的样式，从而改变一个基因的表达。核开关（riboswitch）或 RNA 开关（RNA switch）就是指这一类特殊的 mRNA。

以枯草杆菌胞内参与硫胺素合成和运输的蛋白质为例，其 mRNA 在 5′UTR 含有一段高度保守的 Thi 盒（Thi box），该元件是 TPP 的结合位点。如果胞内的 TPP 水平较高，TPP 就与 Thi 盒结合，诱使 mRNA 提前形成终止子结构，从而迫使转录提前结束；反之，如果胞内的 TPP 不足，就没有 TPP 与 Thi 盒结合，这时 mRNA 形成的是抗终止子结构，转录会继续进行（图 42–11）。

核开关能直接检测到细胞内一些重要的小分子代谢物的水平，并根据细胞的生理需要，打开或关闭相关基因的表达。迄今为止发现的所有核开关都使用高度特异性的适配体结构，来作为相应的目标分子的感应器。一旦目标分子结合，mRNA 在 5′UTR 的二级、三级结构就会发生变化，从而通过影响转录终止、翻译起始或 mRNA 后加工来改变基因的表达。现代的细胞核开关的存在表明，RNA 也能形成可与蛋白质相媲美的复杂结构。而且，核开关省去了使用调节蛋白来调节基因表达，因此更为经济。

（三）抗终止作用

这种手段是通过修饰 RNA 聚合酶，使之忽略终止子结构，从而导致转录发生通读（read-through）。细菌体内有些蛋白质能阻止不依赖于 Rho 因子的终止子结构的形成，在转录的终止阶段来控制基因的表达。以大肠杆菌为例，其 *Bgl* 操纵子编码的蛋白质都与 β- 葡糖苷的吸收和代谢有关。该操纵子受到 β- 葡糖苷的诱导，并受到 BglG 蛋白的正调控。BglG 是作为一种抗终止蛋白起作用的，其作用的机



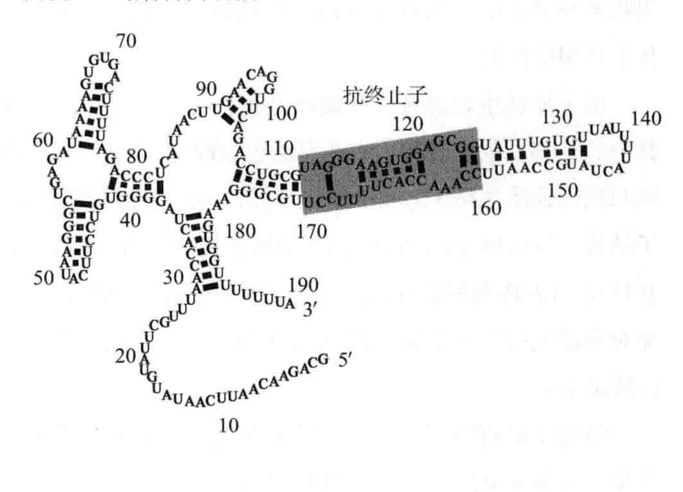


图 42-11 枯草杆菌控制 TPP 合成和运输的核开关的结构及其作用机制

制是与 RNA 转录物上一段特殊的序列结合，阻止终止子结构的形成（图 42-12）。

BglG 有磷酸化和去磷酸化两种形式，但只有磷酸化形式才具有抗终止作用。BglG 的磷酸化由膜上的 BglF 决定。在有 β- 葡萄糖苷的时候，BglF 通过双组分基因表达调控系统（参看本章第四节）激活 BglG，使其磷酸化。在诱导物 β- 葡萄糖苷的作用下，磷酸化的 BglG 形成二聚体，并与 *bgl* 操纵子的 RNA 转录物上的一段特殊序列结合，稳定抗终止子结构，同时阻止终止子形成，从而使转录得以继续下去。

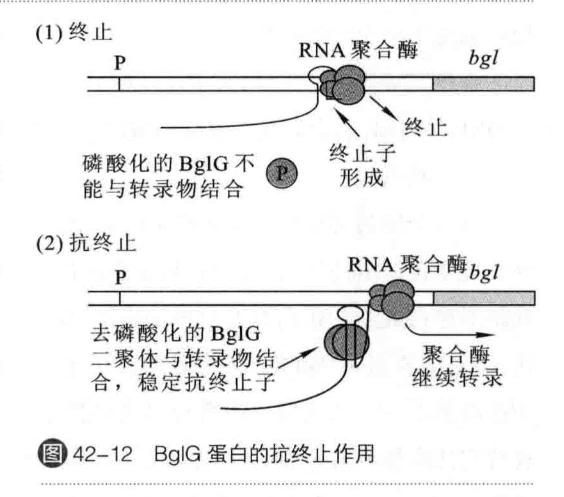


图 42-12 BglG 蛋白的抗终止作用

框 41-1 生化研究动态——探测氟离子的核开关

氟化物是地壳第 13 种最常见的元素，人们很早就开始在牙膏中添加氟化物，这被认为可以减少蛀牙。这种效果在很大程度上是由于氟化物结合到牙齿的釉质上，坚固牙釉质，使其抵抗细菌在口腔内产生的酸性物质。此外，高浓度的氟化物对细菌是有毒的，因此很多人认为，氟化物这种抗菌活性也可以帮助预防蛀牙。那么，细菌面对有毒的氟化物会不会做出某种保护性反应呢？

对于这个问题，耶鲁大学的研究人员给出了肯定的答案。在 2011 年 12 月 22 日的 *Science* 杂志上，最先发现核开关的 Ronald Breaker 教授又报道了一种新的核开关。这种核开关对氟离子是高度特异性的，与氟同族的离子（如 Cl^- 子）结合不上。许多细菌，包括那些导致蛀牙的细菌，能通过这种核开关探测细胞内氟化物的水平，并激活细菌防御。这些核开关的感应器是专门为探测氟化物而设计的。

由于氟离子和检测它的 RNA 核开关都是带负电荷的，因此起初很多人认为它们不大可能有这样的核开关。但是 Breaker 却在许多细菌和古菌体内，发现了超过 2 000 种专门探测氟化物的核开关。这些核开关能激活将氟运出细胞的酶的基因表达，从而抵消氟化物对细菌的毒性作用。

通过追踪和比较存在于很多物种中的氟化物核开关，Breaker 认为这些核开关的历史悠久。数亿年来，许多生物在进化的历史长河中都需要克服氟化物的毒性，因此它们进化了精确的传感器和防卫机制来保护自己。研究和深入了解这些传感器和防御机制，可以让科学家利用这些机制，并有可能在将来制造出更高效对抗蛀牙细菌的氟化物。

第三节　在翻译水平上的调控

翻译水平的调控手段主要有：反义 RNA（antisense RNA）、核开关、自体调控、mRNA 的二级结构和稳定性以及严紧应答等，现分别加以介绍。

一、反义 RNA

反义 RNA 是指与特定目标 RNA 分子（通常是 mRNA）因存在互补序列而发生配对，从而调节目标 RNA 功能的 RNA 分子。

反义 RNA 可参与 DNA 复制和基因转录的调控，而对基因表达的调控主要在翻译水平上，少数在转录水平上。反义 RNA 可能起负调控，也可能起正调控的作用。

（一）反义 RNA 的负调控作用

1. micF-RNA

起负调控作用的反义 RNA 中最典型的代表是 micF-RNA，它由 *micF* 基因编码，是大肠杆菌 OmpF-mRNA 的反义 RNA，由 T. Mizuno 在 1983 年发现。

OmpC 和 OmpF 属于孔蛋白，位于大肠杆菌外膜上，与细胞的渗透压调节有关，分别由分属不同操纵子的 *ompC* 和 *ompF* 基因编码。这两种孔蛋白在外膜上形成的小孔构成了溶质进入细胞的通道，但 OmpF 形成的孔道要大于由 OmpC 形成的孔道，以便让细胞在低渗的环境中加强吸收。

当环境中的渗透压变化时，位于内膜上的 EnvZ 蛋白能够监测到所发生的变化，并通过 OmpR 蛋白调节 OmpC 和 OmpF 的翻译，使得大肠杆菌能够适应环境中的渗透"冲击"。

OmpR 是一种调节蛋白，它有磷酸化和去磷酸化两种形式，但只有磷酸化的形式（OmpR-P）才能与调节位点结合。OmpR-P 能激活或阻遏一个基因的表达，具体情况与它的结合位点与启动子的相对位置有关（图 42-13）。*ompF* 基因是被激活还是被阻遏，分别受其启动子上游的高亲和位点和低亲和位点控制，而 *ompC* 基因和 *micF* 基因的激活由其低亲和位点控制。在高渗的环境中，EnvZ 的 1 个 His 残基自磷酸化，随后，它将磷酸基团转移给 OmpR 的 1 个 Asp 残基，致使细胞内 OmpR-P 的水平提高，这时低亲和位点和高亲和位点都能结合到 OmpR-P，而最终导致 OmpC 表达增加和 OmpF 表达降低。

为了进一步阻止 OmpF 的合成，有一种叫 micF-RNA 的反义 RNA 与 OmpC 同时转录。巧合的是，micF-RNA 与 OmpF-mRNA 的 5′端有部分序列是互补的，因此可以配对形成双链，从而阻断 OmpF 的翻译（图 42-14）；在低渗环境中，去磷酸化的 OmpR 水平提高，而 OmpR-P 水平降低，低水平的 OmpR-P 只能与其高亲和位点结合，这最终导致了 OmpF 表达被激活，OmpC 表达则受阻。

（二）反义 RNA 的正调控作用

DrsA 是在大肠杆菌中发现的又一种反义 RNA，它既可以和 *hns*-mRNA 互补配对，又可以和 *rpoS*-mRNA 互补配对，但对这两种 mRNA 翻译的影响正好相反，前一种配对掩盖了 *hns*-mRNA 在 5′端的核糖体结合位点（RBS），从而导致翻译受阻，后一种配对则暴露出 *rpoS*-mRNA 上的 RBS，从而激活翻译（图 42-15）。

二、核开关

除了可以在转录水平上，核开关还可以在翻译的水平上控制一个基因的表达。以粪肠球菌（*Enterococcus faecalis*）体内控制 SAM 合成的核开关为例，催化 SAM 合成的酶在 mRNA 的 5′UTR 中，除了含有 SD 序列以外，还有一段反 SD 序列。

如图 42-16 所示，如果细胞内的 SAM 不足，就不会有 SAM 的结合，这时 SD 序列不会与反 SD 序列配对，因此催化 SAM 合成的酶照常进行翻译；相反，如果细胞内的 SAM 足够，就会有 SAM 结合，这

图 42-13 渗透压变化改变 OmpF 和 OmpC 表达的机理

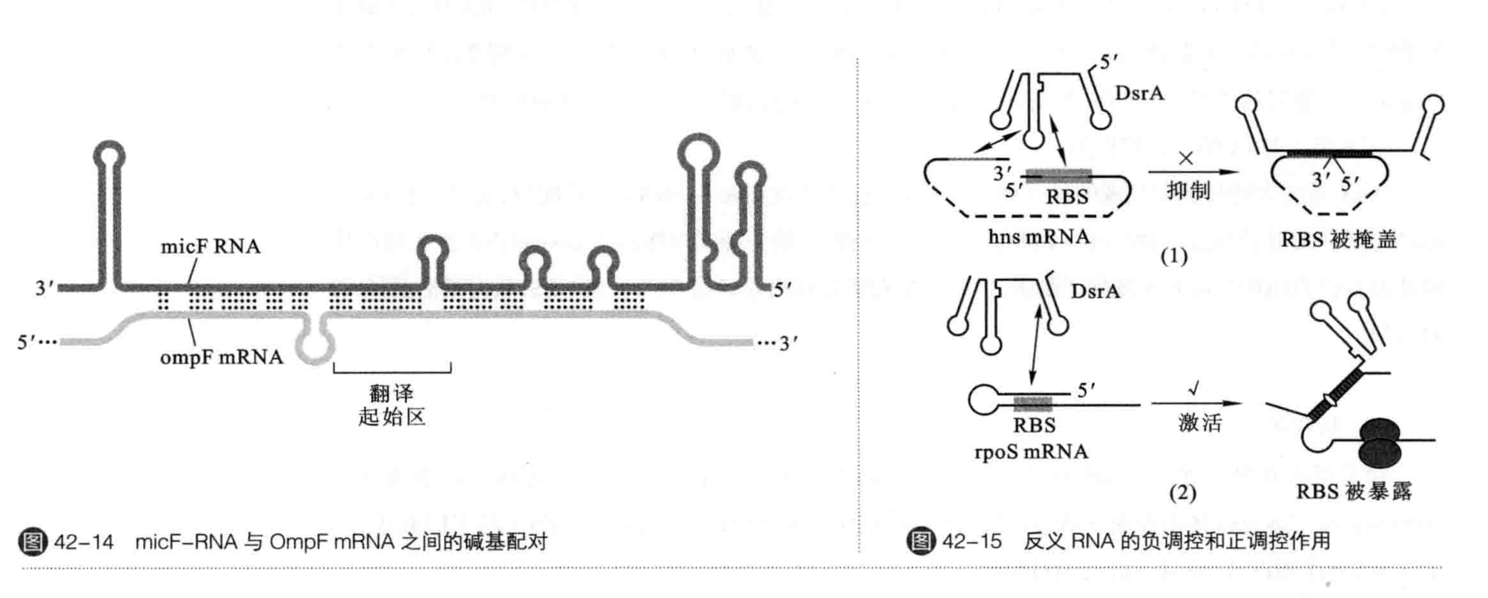


图 42-14 micF-RNA 与 OmpF mRNA 之间的碱基配对

图 42-15 反义 RNA 的负调控和正调控作用

时 mRNA 的构象受到诱导而发生变化，SD 序列与反 SD 序列发生配对，SD 序列将无法被核糖体识别而导致翻译无法启动。

框 42-2　生化新突破——原核生物的 CRISPR 系统

自从在真核生物体内发现 RNA 干扰(RNAi)现象以来,分子生物学家一直在想:原核生物内有没有类似的系统用来对付入侵的病毒。最新的研究显示,原核生物虽然没有 RNAi,但却有另外一个基于 RNA 的防护系统,可以破坏入侵的病毒基因组。事实上,原核生物的这种叫 CRISPR 的系统威力更强大,可同时对付 RNA 病毒和 DNA 病毒,而不像 RNAi 只能对付 RNA 病毒。CRISPR 表示的意思是成簇有规律间隔短回文重复序列(clustered regularly interspaced short palindromic repeats)。

原核生物染色体上的 CRISPR 区域实际上是一种存储外来病毒序列的记忆库。它由许多不同的病毒序列和相同的重复序列交替排列组成。CRISPR 系统提供了对任何含有相同和相近序列的病毒的防护。CRISPR 系统中的蛋白质,即 CRISPR 相关蛋白(CAS 蛋白)具有 2 个功能(图 42-17):一是利用存储的序列信息,去识别入侵的病毒基因组并摧毁它们;二是参与获得和储存病毒的序列片段,这一过程还不清楚。CAS 蛋白由位于 CRISPR 序列上游的基因编码。CRISPR 区作为一个整体先被转录,然后被 CAS 蛋白在每一个重复序列的中间切开,产生单个病毒特异性片段。如果其中的某一个片段可以和外来入侵的病毒 DNA 或 RNA 互补配对,CAS 蛋白就可以将病毒的 DNA 或 RNA 降解。已发现,CRISPR 系统广泛分布在古菌和细菌之中,大约 90% 已测过基因组序列的古菌和 70% 已测过基因组序列的细菌都含有这个系统。

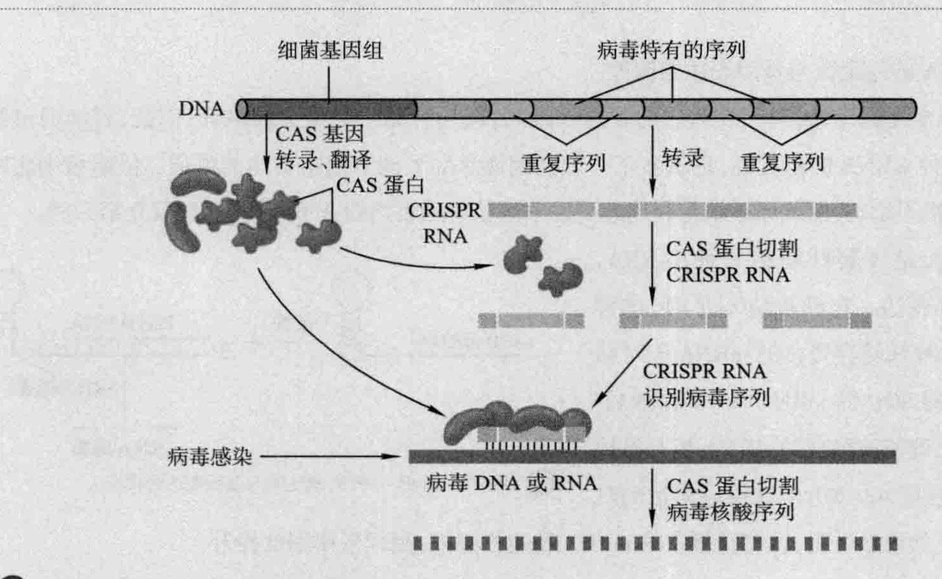


图 42-17　原核生物的 CRISPR 系统(Madigan 等,2010)

现在,科学家已经可以在一种细菌的 CRISPR 序列库中人为地添加其他序列,从而让这种增容过的细菌能对付以前不能对付的外源核酸。例如,在库中添加与水解青霉素的内酰胺酶基因同源的序列,可让一种细菌降解外来带有内酰胺酶基因的质粒 DNA。

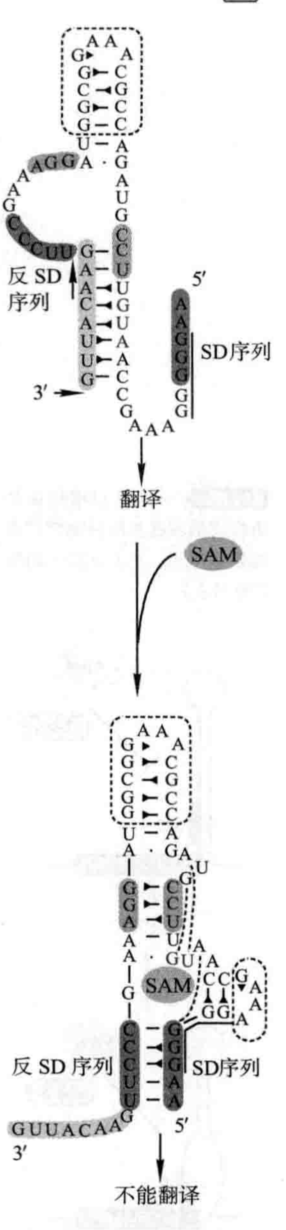


图 42-16　粪肠球菌控制 SAM 合成的核开关的结构及其作用机制

三、自体调控

自体调控(autogenous control)是指一个基因产物对自身的表达产生激活或抑制的现象,这实际上是一种在基因表达水平上的反馈。它既可以在转录水平上(如 AraC 抑制自身的转录,参看“阿拉伯糖操纵子”),也可以在翻译水平上进行。这里以核糖体蛋白的自体调控为例,只介绍后一种情形。

核糖体蛋白的基因组织成多个操纵子,大多数操纵子含有组成大、小亚基的核糖体蛋白的基因,某些与翻译有关的蛋白质的基因也包含在内。在每一个核糖体的操纵子上都有一个基因兼做调节基因,其蛋白质产物能够与自身 mRNA 5′端结合,从而抑制自身的翻译。

以 *s15* 操纵子为例(图 42-18),它含有的结构基因是编码蛋白的 *s15* 基因和编码多聚核苷酸磷酸化酶的 *pnp* 基因。与其他核糖体蛋白一样,S15 在翻译以后就与 rRNA 组装成核糖体。但如果 S15 量多于 16S rRNA,就会与自己的 mRNA5′端结合,抑制自身的翻译,以便和 rRNA 的合成同步。

Quiz5　你认为 S15 与 PNP 受同一个操纵子控制有何益处?

四、mRNA 的二级结构与基因表达的调控

mRNA 的二级结构不仅可以影响到 mRNA 的稳定性，还会影响到 RBS 的可得性。反义 RNA 一节已显示了 mRNA 的 RBS 对于翻译的重要性，下面以单核细胞增生性李斯特菌(*Listeria monocytogenes*)为例，进一步说明 mRNA 二级结构的变化是如何影响 RBS 的可得性，进而调节翻译的。

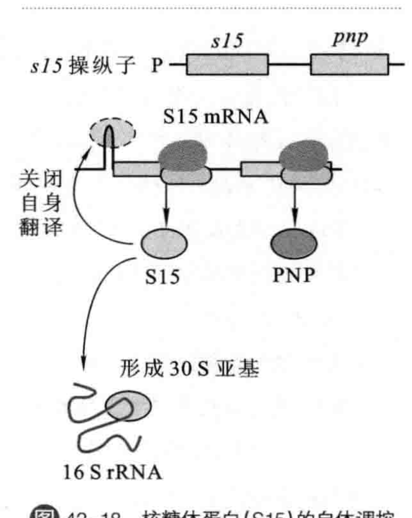


图 42-18　核糖体蛋白(S15)的自体调控

单核细胞增生性李斯特菌是一种可导致人类食物中毒的病原菌，其毒性基因只有在菌体进入宿主体内以后才会表达，而控制毒性基因表达的源头在于温度。PrfA 是一种激活蛋白，负责激活与毒性有关的基因表达。有趣的是，PrfA 在 37℃能表达，在 30℃则不表达，可是它的转录在两种温度下都能进行，看来控制的位点只能是在翻译水平上了。原来是在 30℃或更低的温度下，PrfA-mRNA 上的 SD 序列与其他区域配对形成链内双链，致使 RBS 被掩盖，翻译因此受到抑制；而在 37℃ 下，配对区域热变性，SD 序列暴露，核糖体可以与之结合，翻译便可以进行了(图 42-19)。

Quiz6　一些来自嗜热菌的蛋白质基因在大肠杆菌体内表达效率很低，你认为其中的原因是什么?

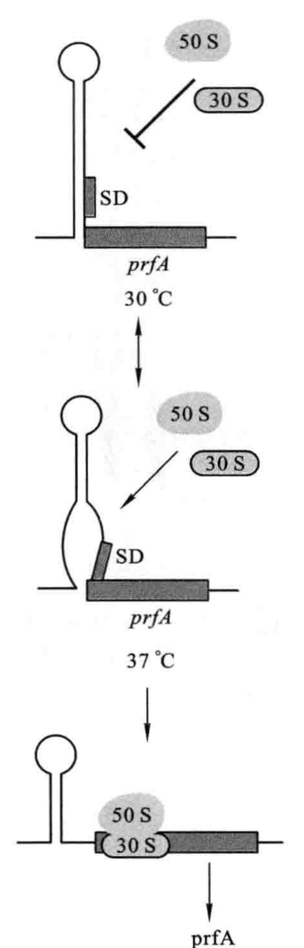


图 42-19　温度对李斯特菌 PrfA 蛋白的表达调控

五、mRNA 的稳定性与基因表达的调控

既然 mRNA 是翻译的模板，那么它的稳定性越高，就会有更多的机会被翻译，因此，若能通过某种手段控制一种 mRNA 的稳定性，也就多了一种在翻译水平上调节基因表达的渠道。但原核细胞的 mRNA 本来就很不稳定，因此原核细胞利用此手段来调控基因表达的并不多见，这里仅介绍一例。

RyhB-RNA 是大肠杆菌的一种小 RNA，仅在铁饥饿时表达。在铁供应不足时，这种 RNA 与胞内 6 种铁储存蛋白的 mRNA 配对结合，包括含铁的 SOD 的 mRNA(sodB-mRNA)，形成双链区域，致使被结合的 mRNA 更容易被 RNA 酶 E 水解(图 42-20)；但铁供应充足的时候，RyhB-RNA 的表达受阻，铁储存蛋白的 mRNA 稳定性提高，翻译效率因此提升。

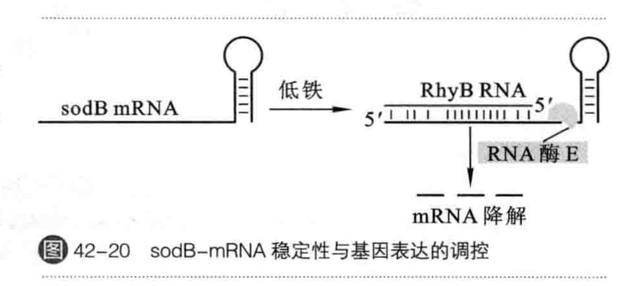


图 42-20　sodB-mRNA 稳定性与基因表达的调控

六、严紧应答

严紧应答专指细菌在氨基酸饥饿的时候，胞内所发生的各种代谢变化。它们主要包括：rRNA 和 tRNA 合成量的急剧下降；mRNA 合成下降；蛋白质降解加强；核苷酸、糖和脂质的合成下降。严紧应答的意义在于使细胞“勒紧裤带”，节省能量，以渡过难关。

大肠杆菌内能够感应到氨基酸饥饿信号的是 RelA 蛋白，它又被称为严紧因子(stringent factor)，具有依赖于核糖体的(p)ppGpp 合成酶活性。此酶活性在氨基酸饥饿的时候被激活，以合成被称为魔斑(magic spot)分子的 pppGpp 或 ppGpp，从而诱导严紧应答。RelA 缺失的大肠杆菌突变体对于氨基酸的饥饿无严紧应答。

严紧应答发生的基本步骤是：在氨基酸饥饿的时候，细胞内空载 tRNA 开始积累，并有机会进入 A 部位。核糖体 50S 亚基上的 L11 蛋白正好位于 A 部位和 P 部位的附近，能够对 A 部位上正确配对的但是空载的 tRNA 做出反应，其构象变化可激活与核糖体结合的 RelA 的酶活性。RelA 受激活后，便开始催化 pppGpp 和少量的 ppGpp 的合成。于是细胞内的 pppGpp 和 ppGpp 迅速积累，在氨基酸饥饿几秒钟以后就达到最高水平。pppGpp 和 ppGpp 的作用对象主要是 rRNA 的操纵子，能抑制 rRNA 基因的转录起始和延伸。而一旦 rRNA 的合成受阻，必然会影响核糖体蛋白的合成，进而影响到其他蛋白质的合成，并最终带来各种后继效应。

一旦氨基酸供应正常，严紧应答便迅速消退，pppGpp 和 ppGpp 被水解。先由 *gpp* 基因的产物将 pppGpp 降解为 ppGpp，再由 SpoT 蛋白将 ppGpp 降解成 GDP。随着 pppGpp 和 ppGpp 的水解，RNA 聚合酶开始转录原来在严紧应答中受到抑制的基因。

Quiz7 有两种突变，一种在氨基酸没有饥饿的情况下也能产生严紧应答，另一种即使氨基酸发生饥饿也不发生严紧应答。你认为这两种突变分别是什么？

第四节　环境信号诱发的基因表达调控

原核生物所生活的环境变幻莫测，因此需要根据环境的变化及时调整特定基因的表达状况。双组分基因表达调控系统（two-component system of gene regulation）可让一种生物能对环境中的各种信号刺激做出有利的反应，故广泛存在于原核生物，但也存在于单细胞真核生物、真菌和高等植物中。

该系统主要由 2 种成分构成（图 42-21 和图 42-22）：①第一种负责感应环境中的信号，称为感应子激酶（sensor kinase）。它有 2 个结构域，一个直接感应信号，称为感应子（sensor），另一个为自激酶（autokinase），在感应子接受信号以后，它的 1 个 His 残基发生磷酸化。②第二种负责调节细胞做出什么样的反应，也有 2 个结构域，一个称为接受子（receiver），其内的 1 个 Asp 残基可以从第一种成分中的 His 残基接受磷酸基团，另一个结构域负责与 DNA 结合。当第二种成分从第一种成分接受磷酸基因以后，就与 DNA 结合，从而调节一系列基因的表达。

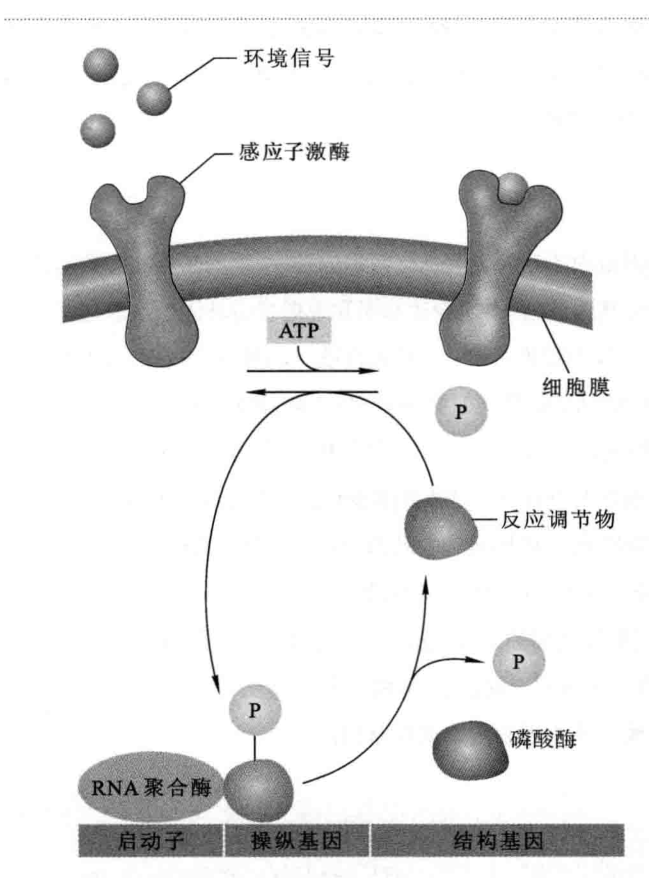


图 42-21　双组分基因表达调控系统图解

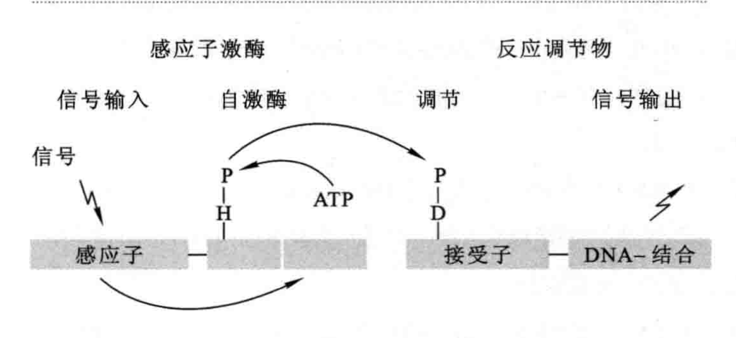


图 42-22　双组分基因表达调控系统中的两种成分的结构与功能

前面提到的 BglF 蛋白和 BglG 蛋白，以及 EnvZ 和 OmpR 就都分别属于双组分基因表达调控系统中的第一、二种成分。

某些细菌除了能对环境中的非生物因素的变化做出反应以外，还能对环境中同类生物的存在做出反应。群体感应（quorum sensing）就属于这类反应，它是细菌根据细胞密度变化进行基因表达调控的一种生理行为。当细菌密度达到临界阈值时，细菌能释放一些特定的自诱导信号分子，从而调节本种群或同环境中其他种群的群体行为。

e42-1 群体感应

第五节　在基因组水平上的全局调控

以上集中介绍了单个基因或者一组基因(操纵子)的局部表达调控机制,并没有涉及整个基因组的全局表达调控。然而,对任何一种生物来说,其基因组内的各个基因的表达并不是“各自为政”的,而是作为一个整体受到控制。这种全局性的调控对一种生物的生存和繁殖更为重要。下面就结合几种机制进行简单的介绍。

一、调节子

调节子(regulon)即调谐子,是指不同的操纵子受同一种调节蛋白调节的网络结构,通过调节子可以协调一系列在空间上分离但功能上相关的操纵子的转录活动。

大肠杆菌有 2 个重要的调节子最具代表性,一个是 *sos* 调节子,它负责协调参与 SOS 修复有关的多个操纵子基因的表达,起协调作用的调节蛋白是 LexA 蛋白(参看第三十四章“DNA 损伤、修复和突变”);第二个调节子与 CAP-cAMP 有关,它参与协调与碳源分解代谢有关的多个操纵子近 200 多个基因的表达。除此以外,大肠杆菌还有与营养状态变化反应、氧饥饿反应、氧化还原状态变化反应、酸碱反应、热激反应和冷激反应等有关的各种调节子结构。

二、噬菌体基因表达的时序控制

一种噬菌体在感染宿主细胞以后,其基因组内的基因表达具有明确的顺序,某些基因在噬菌体进入宿主细胞后就开始表达,属于所谓的早期表达基因,然后中期表达基因和晚期表达基因依次表达。那么基因表达的时序性是如何实现的呢?对于这个问题很难给出一个明确而统一的答案,这是因为不同的基因组使用的控制策略并不总是相同的。有的比较简单,像前面提到的 SPO1 噬菌体是使用 σ 因子的级联来表达不同时期的基因。而 T7 噬菌体使用的策略是基于不同 RNA 聚合酶之间的转换,其早期基因转录使用宿主细胞的 RNA 聚合酶,但在转录的早期基因中有一个是噬菌体编码的高度特异性的 RNA 聚合酶,由这一种 RNA 聚合酶催化中期和晚期基因的转录,而中期表达的基因产物之一是能够中和宿主 RNA 聚合酶的活性,以保证晚期胞内的转录活动完全是噬菌体特有的。有的则比较复杂,如 λ 噬菌体。

λ 噬菌体的生长不仅涉及在溶原期和裂解期之间的选择,还涉及在溶原期或裂解期内不同的生长阶段,即从早期、早前期、中期到晚期的过渡。在每一个阶段,λ 噬菌体所表达的基因不尽相同。为了保证基因表达的高度时序性,λ 噬菌体一共使用了 5 种手段来调控基因的表达。

e42-2 λ 噬菌体基因表达的时序性调控

科学故事——“核开关”的发现

RNA 似乎总是给我们带来惊喜,这次的惊喜是发现了核开关,发现人是耶鲁大学的 Ronald Breaker。

虽然早在 20 世纪 60 年代就有人意识到有可能存在核开关,但并没有人真正去做有关的研究。即便是到了 90 年代,绝大多数生物学家还是认为行使生物感应功能是蛋白质的“特权”,没有 RNA 的份。

在 80 年代核酶被发现以后,科学家认为,在 30 亿 ~40 亿年之前,即在 DNA 或蛋白质出现之前,有一个 RNA 世界,在 RNA 世界里,由 RNA 说了算!

90 年代后期,Breaker 设想,如果确实有 RNA 世界,那过去就一定存在 RNA 开关来检测代谢状态,以便在瞬间打开和关闭某些代谢过程。而且,在现在的活细胞内极有可能残留有天然的核开关。随后他在体外做了一些实验,结果很容易地得到了几种人造的核开关。

在 2000 年夏天的实验室一次晨会上,Breaker 中途停止了会议,他对自己的学生说,既然在试管里很容易制造出核开关,而且运作得很好,它们一定也存在于现代的细胞里。于是,他和他的弟子打了两个赌,第一个赌是如果现代细胞里有核开关,它们就一定位于 mRNA 上,因为 mRNA 能够控制蛋白质的合成;第二个赌是生物学家一定曾经遇到过核开关,只是没有意识到,因为每一个人都期待感应代谢物水平变化

的因子应该是蛋白质。假如一个研究小组发现一个基因的表达受代谢物 X 控制，代谢物 X 含量升高，基因就会关闭，但是，研究者却苦于找不到那种结合 X 并关闭基因的蛋白质。

就在那次会议的第二天，他的一个博士后在图书馆里好不容易翻阅到一篇有关维生素 B_{12} 生物感应器的论文，那没有找到的感应器蛋白事实上就是一种核开关。随后，Breaker 的研究小组进行了大量的实验，证明了大肠杆菌细胞内存在维生素 B_{12} 的核开关，并且发现这个核开关在结合 B_{12} 以后会改变形状，而它与 B_{12} 结合也是高度特异性的，即便是它的类似物存在也不会引起类似的反应。

在随后更多的文献调研中，他们发现并证实了其他 5 种类似的被遗漏的核开关，但他们并没有就此罢休，认为还应该有更多的核开关会被发现，为此他们建立了一种快速的在微生物基因组里寻找类似的核开关生物信息学手段，很快他们就有了惊人的发现！

在 2004 年 2 月召开的美国高级科学促进会年会上，Breaker 报道了在枯草杆菌内他发现的第 8 个细菌核开关。这个核开关不同于早先发现的 7 个核开关，是一种核酶，即是用核酶来调节 6- 磷酸葡糖胺的水平。6- 磷酸葡糖胺是细菌细胞壁合成的关键成分，在转氨酶催化下，6- 磷酸果糖和谷氨酰胺之间发生转氨反应，生成 6- 磷酸葡糖胺和谷氨酸。当细胞内的 6- 磷酸葡糖胺浓度较高时，它就与转氨酶的 mRNA 结合，激活了它的核酶活性，而使 mRNA 发生自我切割，从而关闭转氨酶的合成。奇怪的是，mRNA 被剪切的部位并非蛋白编码区，但却仍然能成功破坏基因表达。

目前研究发现位于 mRNA 非编码区的核开关广泛存在于细菌中。估计枯草杆菌约有 2% 基因的表达受到这种机制的调控。比较典型的几种是分别由维生素 B_{12}、TPP、FMN、SAM、鸟嘌呤、腺嘌呤、6- 磷酸葡糖胺、甘氨酸和赖氨酸控制的核开关。

核开关的发现不仅为 RNA 俱乐部补充了新的成员，还为 RNA 世界的假说提供了新的证据。目前 Breaker 不仅在寻找新的核开关，还在尝试使用人造的核开关来杀死细菌和肿瘤细胞。随着越来越多的复杂 RNA 分子被发现，人们越来越清楚地认识到 RNA 世界不仅存在于遥远的过去，实际上，我们今天就生活在 RNA 世界。

e42-3 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Regulation_of_gene_expression（维基百科有关基因表达调控的内容）
2. http://themedicalbiochemistrypage.org/gene-regulation.html（完全免费的医学生物化学课程网站，有关基因表达调控的内容）
3. http://www-biology.ucsd.edu/classes/bimm100.WI00/XIV.Lambda.html（美国加利福尼亚大学圣地亚哥分校一个生物学网站，有关 λ 噬菌体的内容）
4. http://www.dartmouth.edu/~cbbc/courses/bio4/bio4-lectures/ProkGeneControl.html（美国达特茅斯学院一个生物学网站，有关原核基因表达调控的内容）

参考文献：

1. White M F, et al. Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Molecular Cell*, 45: 1-11, 2012.
2. Karginov F V, Hannon G J. The CRISPR system: small RNA guided defense in bacteria and archaea. *Mol Cell*, 37: 7-19, 2010.
3. Walter N G, et al. Trans-acting glmS catalytic riboswitch: Locked and loaded. *RNA*, 13: 468-477, 2007.
4. Mandal M, Breaker R R. Gene regulation by riboswitches. *Nat Rev Mol Cell Biol*, 5: 451-63, 2004.
5. Adhya S, et al. Transcriptional regulation by cAMP and its receptor protein. *Annu Rev Biochem*, 62: 749-795, 1993.
6. Johnson R C. Mechanism of site-specific DNA inversion in bacteria. *Curr Opin Genet Dev*, 1: 404-411, 1991.
7. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol*, 3: 318-356, 1961.

第四十三章 真核生物的基因表达调控

与原核生物一样，真核生物的基因表达也受到严格的调控，例如，编码血红蛋白 2 个亚基的基因只在红细胞中表达，人肝细胞在胚胎期合成甲胎蛋白，成年后就很少合成了，然而，当肝细胞癌变以后，编码甲胎蛋白的基因又会开放。一个典型的高等真核生物细胞所表达的基因通常只占基因总量的 10%~20%。

真核生物与原核生物不仅在基因组的结构、基因表达的具体步骤上存在许多差异，在基因表达的调控方面也具有很多显著的差别，这些差别主要包括以下几点：

(1) 调控的目的　原核生物基因表达调节的目的是为了更有效和更经济地对环境的变化做出反应，而多细胞真核生物基因表达调节的主要目的是细胞分化，这需要在不同的生长时期和不同的发育阶段具有不同的基因表达样式。

(2) 调控的层次　原核生物基因表达调控主要集中在转录水平，而真核生物基因表达的转录后水平调节与其在转录水平上的调节同样重要，其中某些调控层次是真核生物特有的，比如染色质水平、RNA 后加工水平（如可变加尾、可变剪接和编辑等）和 RNA 干扰等。

Quiz1 为什么真核生物很少使用操纵子这种结构来调节基因表达？

(3) 调控的手段　真核生物主要使用正调控，而原核生物主要使用负调控。另外，原核生物绝大多数的基因组织成操纵子，而真核生物一般无操纵子结构。虽然秀丽线虫的基因组约 15% 的基因组成类似于原核生物的操纵子结构，但线虫操纵子转录出来的共转录物需要被切割成单独的 mRNA，并在 5′ 端加上帽子，只有这样的 mRNA 才能被翻译。

真核生物基因表达调控不仅可以使细胞更好地适应内外环境的变化，避免资源和能源的浪费，还可以提高基因的编码能力，让一个基因编码出多种蛋白质。此外，这也是多细胞真核生物细胞分裂、分化、癌变、衰老和死亡的分子基础。

虽然真核生物的基因表达调控机制较为复杂，但所有的调控都可以归为两类：一是短期调控（short-term regulation）——见效快，持续时间短。通过这种机制，基因可根据环境的变化和细胞的需要迅速地开放或关闭。二是长期调控（long-term regulation）——见效慢，持续的时间久。这种机制主要应用于生物发育和细胞分化。

本章将从几个不同的层次分别介绍真核生物的基因表达的调控机制，并将重点介绍真核生物几种特有的调控模式。

第一节 在染色质水平上的调控

原核生物和真核生物在 DNA 转录及其调控方面一个显著的差别就是 DNA 模板的状态不同。以细菌为例，其染色体 DNA 绝大多数几乎处于完全裸露的状态，而真核生物的细胞核 DNA 大部分被结合的组蛋白遮挡。因此，原核生物的基因表达的“默认状态”是开放，其调控机制主要是通过阻遏蛋白进行的负调控，而真核生物基因表达的“默认状态”是关闭，其调控机制主要是通过激活蛋白进行的正调控。

真核生物的 DNA 与组蛋白以及一些非组蛋白共同构成染色质结构。染色质又可以分为常染色质和异染色质，前者在结构上较为松散，对 DNA 酶的消化比较敏感，而后者在结构上更为紧密，能够抵抗 DNA 酶的消化。实验证明，具有转录活性的区域属于常染色质。

染色质是一种动态可变的结构，其结构的变化能直接影响到基因的表达。影响染色质结构的主要因素有3个：一是组蛋白的共价修饰；二是染色质重塑因子对染色质的作用；三是组蛋白变体(histone variant)取代标准组蛋白。这三个因素都会影响到基因的表达，而且它们之间存在关联。

一、组蛋白的共价修饰对基因表达的影响

已有众多证据表明，真核生物一个核基因在表达前后，其所在位置的染色质结构一般会发生重塑。由于染色质的组成单位是核小体，因此，染色质结构的改变是从核小体的变化开始的，而核小体的变化又是从组蛋白的共价修饰或去修饰开始的。在这里，组蛋白的共价修饰并没有影响一个基因的碱基序列，但却影响基因的表达，进而影响生物的表型，因此这样的遗传方式属于表观遗传(epigenetic)。

Quiz2 真核生物还有哪些机制可产生表观遗传？

► 表43-1 组蛋白不同共价修饰对基因表达的影响以及其他功能

修饰形式	修饰位点	功能
乙酰化	H3K9，H3K14，H2BK5，H2BK20	转录激活
单甲基化	H3K4，H3K5，H3K27，H3K79，H3R17，H4R3，H4K20或H2BK5	转录激活
	H3R8	转录阻遏
双甲基化	H1K26，H3K5，H3K27	转录阻遏
	H3K79	转录激活
三甲基化	H3K4	转录激活
	H3K9，H3K27，H2BK5	转录阻遏
磷酸化	H1S27	转录激活、染色质去浓缩
	H3T3，H3S10，H4S1	有丝分裂、转录激活、DNA修复
小泛素相关修饰物修饰	酵母的H2BK6或H2BK7	转录阻遏
泛酰化	酵母H2BK123	转录激活
	哺乳动物H2B120	减数分裂

组蛋白能够经历的共价修饰包括：乙酰化、甲基化、磷酸化、生物素化、泛酰化、小泛素相关修饰物修饰和ADP-核糖基化等，其中乙酰化对染色质结构和基因表达的影响最大(表43-1)。在一个特定基因附近的组蛋白分子上发生的共价修饰，决定了一个基因能否表达以及表达的强度，于是，一种特定的修饰样式相当于一个控制特定基因表达的密码，即组蛋白密码(histone code)。那些能够改变组蛋白共价修饰模式的酶均是组蛋白密码的调节者。

每一个组蛋白都有一个柔性的N端尾巴，此尾巴伸展在组蛋白核心和DNA双螺旋之外，作为各种修饰酶作用的主要位点，其中的Lys残基为乙酰化位点。催化组蛋白发生乙酰化和去乙酰化反应的酶分别是组蛋白乙酰转移酶(HAT)和组蛋白去乙酰酶(HDAC)(图43-1)。

组蛋白的乙酰化修饰至少具有3个功能：①中和Lys残基上的正电荷而减弱组蛋白与DNA的亲和力；②招募其他刺激转录的激活蛋白和辅激活蛋白；③启动染色质重塑。以上三个方面均有利于基因转录的发生。

属于原核生物的古菌虽然很多也具有组蛋白，并形成核小体，但组蛋白在两端缺乏从核小体伸出的尾巴，所以不像真核生物那样会发生各种共价修饰。

二、染色质重塑对基因表达的影响

染色质重塑是指在依赖于ATP的核小体重塑复合物的调节下，组蛋白受ATP水解的驱动，沿着DNA发生移位，从而导致核小体之间的间距发生改变，甚至产生无核小体染色质或浓缩染色质的过程。

通过重塑，染色质结构可以朝两个相反的方向转变，一种有利于基因的转录，一种则抑制基因的转录。无论向哪一个方向转变，都需要两类蛋白质复合物的参与。如果是向有利于转录的方向进行

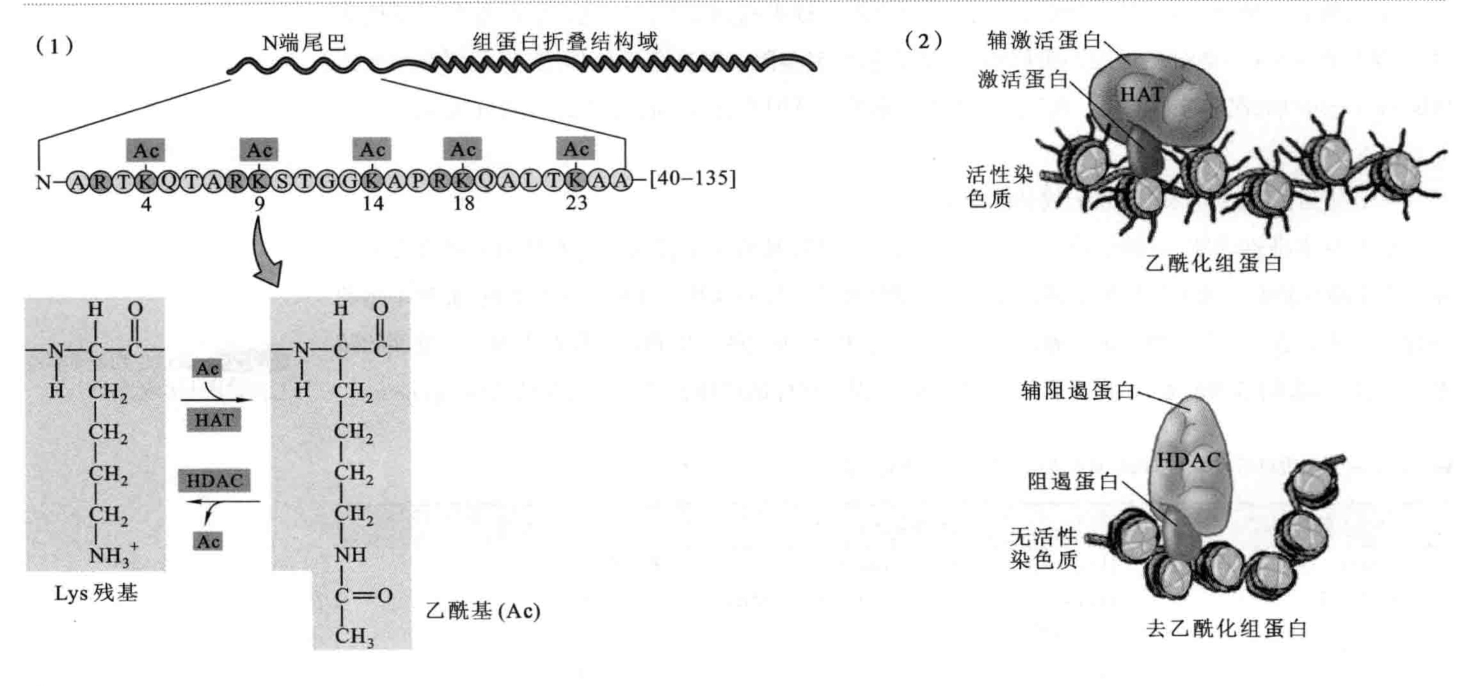


图 43-1 组蛋白乙酰化或去乙酰化与染色质转录活性的关系

重塑，就需要 HAT 复合物和促进染色质采取松散构象的重塑因子。HAT 复合物包括 p300/CBP 复合物和 SAGA 复合物等，其中 p300/CBP 可能是最重要的，因为它可以和一系列参与转录的转录因子相互作用。促进染色质构象松散的重塑因子包括 SWI/SNF 因子和 RSC 复合物的表达。

p300/CBP 复合物由 p300 和 CBP 组成，两者具有相似的结构，其中 CBP 就是 CREB 结合蛋白。p300/CBP 以 3 种方式提高基因的表达：①通过自带的 HAT 活性改变染色质的构象，使其变得松散；②有助于 RNA 聚合酶Ⅱ和基础转录因子招募到启动子上；③充当其他调节转录的分子与基础转录复合物作用的接头分子。SAGA 复合物也是一类带有 HAT 活性的多亚基复合物，同时还具有脱泛酰酶的活性，它主要催化 H3 的乙酰化。

SWI/SNF 因子和 RSC 复合物都是由多个亚基组成，具有依赖于 DNA 的 ATP 酶活性，利用水解 ATP 释放出的能量，使核小体 DNA 在核小体的表面形成突环结构，或者让核小体发生滑移，又或者让组蛋白八聚体转移出去，从而使染色质的结构得以重塑。经过重塑的染色质内的与基因转录有关的各种顺式元件得以暴露，这有利于相应的反式作用因子的识别和结合，从而启动转录。参与重塑过程的重塑因子通常都含布罗莫结构域(bromodomain)，此结构域负责与组蛋白分子上乙酰化的 Lys 残基结合。

如果是向不利于转录的方向进行重塑，就需要 HDAC 复合物和促进染色质浓缩的重塑因子。已发现的SIN3-RPD3复合物就是一种HDAC复合物，其中SIN3是辅阻遏蛋白，RPD3具有HDAC的活性。

在一个细胞内，会同时存在多个不同的染色质重塑复合物(图 43-2)：某些用来松散染色质，使异染色质变成常染色质，从而启动一个基因的表达；还有一些用来浓缩染色质，使常染色质变成异染色质，从而关闭一个基因的表达。总之，如果要激活一个基因的表达，首先就需要一种激活蛋白与该基因附近的增强子结合。一旦激活蛋白结合上去，它一方面将 HAT 复合物(p300/CBP 复合物或 SAGA 复合物等)招募进来，另一方面又通过辅激活蛋白、介导蛋白与基础转录因子、RNA 聚合酶相互作用。招募进来的 HAT 复合物催化附近的组蛋白发生乙酰化修饰，而被修饰的组蛋白又将染色质重塑因子(SWI/SNF 因子或和 RSC 复合物)招募进来。新“入盟”的重塑因子以水解 ATP 为动力，促使染色质的构象变得更加松散，以暴露启动子和其他顺式作用元件，方便基础转录因子和 RNA 聚合酶的识别和结合。一旦基础转录因子和 RNA 聚合酶结合上来，一种庞大的预起始转录复合物便形成了。反之，如果要阻遏一个基因的表达，首先就需要一种阻遏蛋白与该基因附近的沉默子结合。一旦阻遏蛋白结

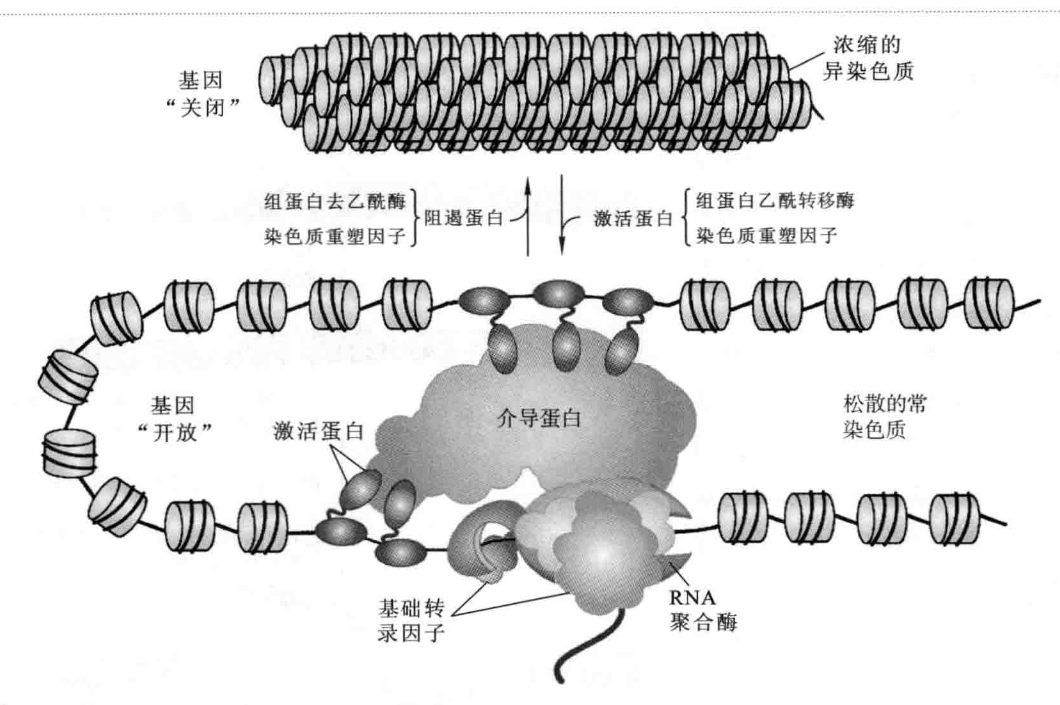


图 43-2 组蛋白的修饰与染色质构象变化的关系(Passarge,2007)

合上去,它会将 HDAC 复合物(SIN3/RPD3 复合物)招募进来。招募进来的 HDAC 复合物催化附近的组蛋白发生去乙酰化反应,而丢掉乙酰基的组蛋白再将促进染色质浓缩的重塑因子招募进来,促使染色质的浓缩,使启动子和其他顺式作用元件难以识别,基因表达之门也就被关闭了。

以酵母细胞参与交配的 HO 内切酶的基因转录激活为例(图 43-3),首先需要一种激活蛋白 SWI5P 与其增强子结合,然后再通过组蛋白乙酰转移酶(SAGA 复合物)、染色质重塑因子(SWI/SNF 复合物)、激活蛋白 SBP 和介导蛋白的依次作用,最后将基础转录因子和 RNA 聚合酶招募到 HO 基因的核心启动子上,由此启动 HO 基因的转录。

三、组蛋白变体对基因表达的影响

所有真核细胞内的组蛋白都具有 5 种标准的形式。这 5 种标准的形式由多个拷贝的基因编码,在进化上高度保守。与此同时,细胞内还存在与这 5 种标准形式相对应的不同变体。这些变体的氨基酸序列与相应的标准组蛋白在一级结构上差别可能很小,也可能很大,表达的方式可能是组成型,也可能具有组织特异性。编码这些变体的基因与相应的标准组蛋白的基因并不呈等位基因的关系,且通常是单拷贝。有些变体与相应的标准组蛋白相比,在与 DNA 结合的性质上差别很大,故它们在取代相应的标准组蛋白以后,会改变核小体和染色质的结构,进而有可能影响到基因的表达。

如果一种变体代替了相应的标准组蛋白参入到一个或多个核小体结构之中,就等于在染色质上创造了一些“特区”,或者在染色质上打上了特殊的标记。这些特化的区域或者特殊的标记可能具有多种不同的功能,但最重要的功能可能是调节特定的基因表达。例如,H2A 有一种最常见的变体叫 H2A.Z,它在 C 端的序列和长度明显不同于标准的 H2A。已发现它在染色质许多特定的区域取代 H2A,而这种取代是目前所有进行测试动物的发育和生存所必需的。

组蛋白变体对基因表达的影响是多方面的。最早在四膜虫细胞内发现,H2A.Z 的参入与具有转录活性的染色质相偶联。后来在酵母细胞中发现,H2A.Z 位于异染色质的两侧,其存在能阻止异染色质向周围常染色质区域扩散。另外,在酵母的常染色质区,几乎所有基因的启动子区都存在带有 H2A.Z 的核小体,它经常出现在无核小体启动子两侧的 2 个核小体内。不过,这种核小体的存在,可能激活也可能阻遏基因的表达。

四、性染色质构象的变化与两性基因剂量之间的平衡

雄性和雌性动物在性染色质的结构上有差别，为了维持两性在基因剂量上的平衡，性染色质的结构在特定阶段发生变化，以激活或抑制整条染色质上的基因表达。例如，雌性哺乳动物在胚胎发育到某个阶段，会随机失活两条X染色质的一条，使之变成高度浓缩的巴氏小体(Barr body)，以维持与只含一条X染色质的雄性哺乳动物具有相同的基因表达水平；在雄性果蝇细胞内，其单个X染色质的转录活性加倍，以维持与雌性果蝇相同的基因表达水平。雄性果蝇细胞内的单个X染色质的转录活性加倍与2种ncRNA——roX1和roX2有关。这2种RNA只在雄果蝇体内表达，功能是负责雄性特异性致死(male specific lethal，MSL)蛋白质复合物的组装。MSL复合物被组装好以后，与X染色质结合，促进X染色质上的H4组蛋白的Lys16(H4K16)发生乙酰化修饰，从而提高X染色质上的基因转录水平。MSL复合物在果蝇基因组上约有35个进入点(entry site)，其中有2个为编码roX1/roX2基因的位点。在雌雄同体线虫细胞(XX)内，一种叫剂量补偿复合物(dosage compensation complex，DCC)的蛋白质与两条X染色质结合，会导致转录活性各自减半，以维持与雄性线虫(XO)相同的基因表达水平。

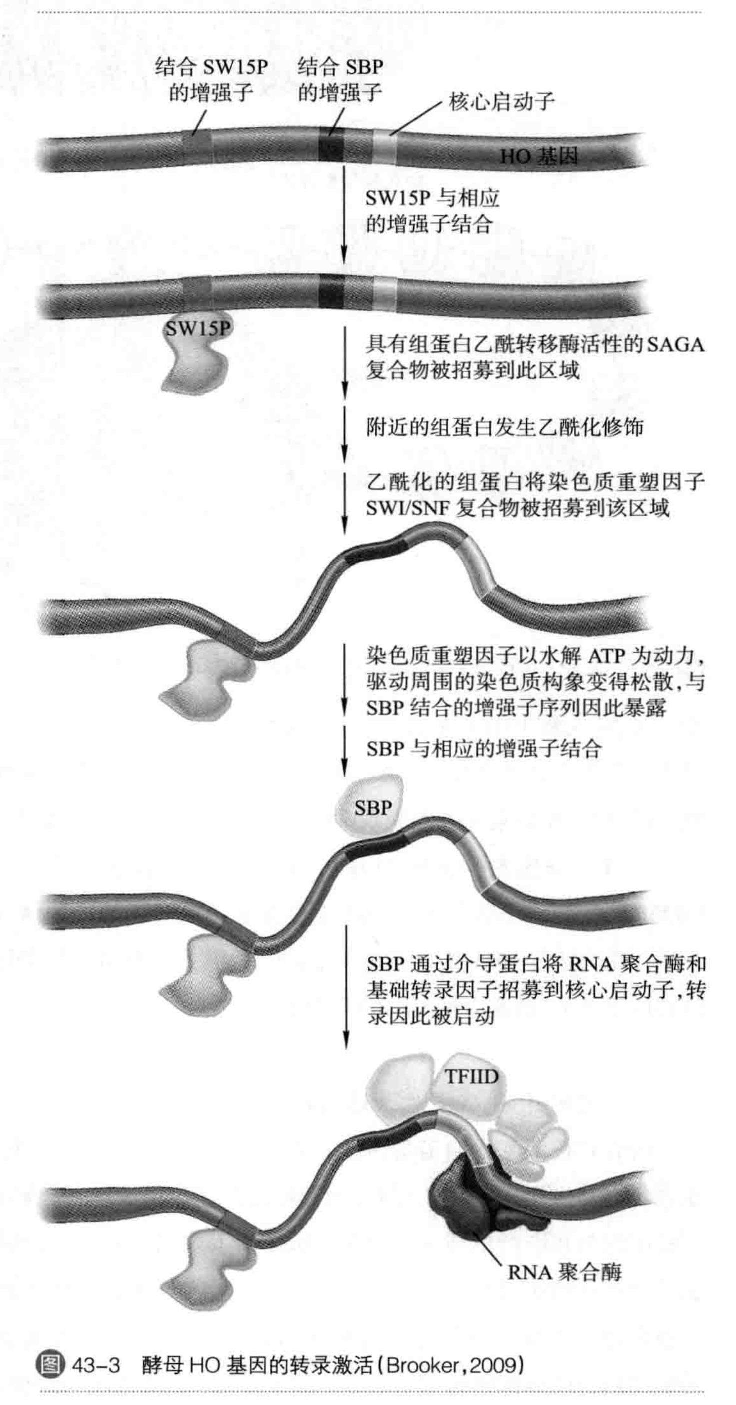


图43-3 酵母HO基因的转录激活(Brooker，2009)

第二节 在DNA水平上的调控

真核生物在DNA水平上调控基因表达的手段有：DNA扩增、DNA重排、DNA甲基化、DNA印记和启动子的可变使用等。

一、DNA扩增

DNA扩增(DNA amplification)是通过增加基因的拷贝数来提高基因表达效率的一种手段，使用这种手段来进行调控的通常是细胞在较短的时间内或在特定的发育阶段需要大量表达的基因。当其他

提高基因表达的调控手段达到极限的时候，DNA 扩增就显得尤为重要。

例如，两栖动物的成熟卵细胞在受精后，其编码 rRNA 的基因(5S rRNA 基因除外)即 rDNA 使用滚环复制大量扩增，拷贝数可增至 2 000 倍，很显然，这是为受精卵在随后分裂和分化过程中需要有大量核糖体来合成大量蛋白质而准备的。再如，氨甲基蝶呤是哺乳动物细胞 DHFR 的抑制剂，如果将哺乳动物细胞放在含有这种抑制剂的培养基上培养，绝大多数细胞会死亡，但少数细胞却能生存下来。这些存活下来的细胞在包含 *DHFR* 基因的 DNA 区段，通过局部的滚环复制扩增了近 4 000 倍，使 DHFR 的表达量显著增加，从而对氨甲基蝶呤产生了抗性。

二、DNA 重排

真核生物利用 DNA 重排来调控基因表达的典型例子，是高等动物 B 淋巴细胞在成熟过程中，编码抗体轻链和重链的基因所经历的程序性重排(programmed rearrangement)。

抗体是动物体液免疫的基础，其近乎天文数字的多样性曾叫人百思不得其解。以人为例，估计 1 个人至少可产生 10^9 种特异性不同的抗体，但人的蛋白质基因总数在 2 万多个。既然蛋白质是由基因编码的，那么这么少的基因如何能够编码出如此繁多的蛋白质呢？

为了解释抗体的多样性，先后有人提出了胚系学说(germ line theory)和体细胞突变学说(somatic mutation theory)。前一种学说认为，一个物种的每一个个体本来就有编码所有抗体的基因，它们都是从亲代遗传而来的；后一种学说认为，胚系细胞含有有限的抗体基因，然而，在 B 细胞成熟的过程中，抗体基因发生了突变，而不同的 B 细胞发生不同的突变，从而产生不同的抗体。这两种学说虽然都含有部分正确的成分，但都不能从根本上解决问题。

就在发现构成抗体的 H 链和 L 链都是由 V 区和 C 区组成以后不久，Dryer 和 Bennett 于 1965 年，提出了一种能节省抗体基因空间的双基因学说(two-gene theory)。该学说认为，一个生物只有一个编码 H 链或 L 链 C 区的基因，但含有多个编码 V 区的基因，既然 C 区至少占据了整个抗体的一半，那么通过这样的组织，首先就减少了至少 50% 的抗体基因。随后，在 B 细胞成熟过程中，可能是 V 区基因和 C 区基因，或者 V 区 mRNA 和 C 区 mRNA，或者 V 区肽段和 C 区肽段再连接起来。由于这种学说与当时传统的基因结构和蛋白质合成的观念相悖，受到了人们的排斥。后来过了 11 年，在限制性内切酶被发现以后，两位在瑞士工作的日本科学家 Nobumichi Hozumi 和 Susumu Tonegawa 用实验证明，Dryer 和 Bennett 的观点基本是正确的。而且，他们还发现了第三个基因片段 V_J，此片段负责将 V_L 与 C 区连接起来。

e32-4 Hozumi 和 Tonegawa 的经典实验

至于重链基因的结构，可使用类似的方法进行确定。不过与轻链基因结构不同的是：重链的基因多了一个多样性片段(diversity，D_H)。

图 43-4 为胚系细胞轻链基因的结构，以及 B 淋巴细胞在成熟过程中所经历的重排反应。从图中可以看出：轻链基因共分为 3 个部分，以 κ 轻链为例，分别是 $L_κV_κ$、$J_κ$ 和 $C_κ$。$J_κ$ 紧靠 $C_κ$，但远离于 $L_κV_κ$。$L_κ$ 编码的是轻链的信号肽序列，它与 $V_κ$ 之间有内含子。$L_κV_κ$ 和 $J_κ$ 具有多个不同的拷贝，以 $L_κV_κ$ 变化最大。

以人的 κ 轻链为例，其基因位于第 2 号染色体上，由 40 个 V 区基因(编码 N 端 95 个氨基酸残基)和 5 个 J 基因(编码另外 13 个氨基酸残基)和 1 个 C 基因组成，其中 J 基因和 C 基因之间是 1 个大的内含子序列。

人的重链基因则位于第 14 号染色体上，约有 50 个 V 区基因、20 个 D 基因，6 个 J 基因和若干个含有内含子的 C 区基因(图 43-5)。由于多了 D 基因，故能创造更多的多样性。

如果是重链，就需要发生 2 次重排反应，第 1 次是在 D-J 之间，第 2 次在 V-DJ 之间。与轻链重排反应不同的是：在 V 基因、D 基因或 J 基因被切开以后，细胞内的末端脱氧核苷酸转移酶(terminal deoxynucleotidyl transferase，TdT)会在被切开的 3′ 端随机添加若干个脱氧核苷酸(最多可达 15 个)，从而进一步扩大了抗体基因的多样性。

Quiz3 如果克隆羊多莉是由 B 淋巴细胞或 T 淋巴细胞克隆得到的，你认为多莉羊会有什么异常？

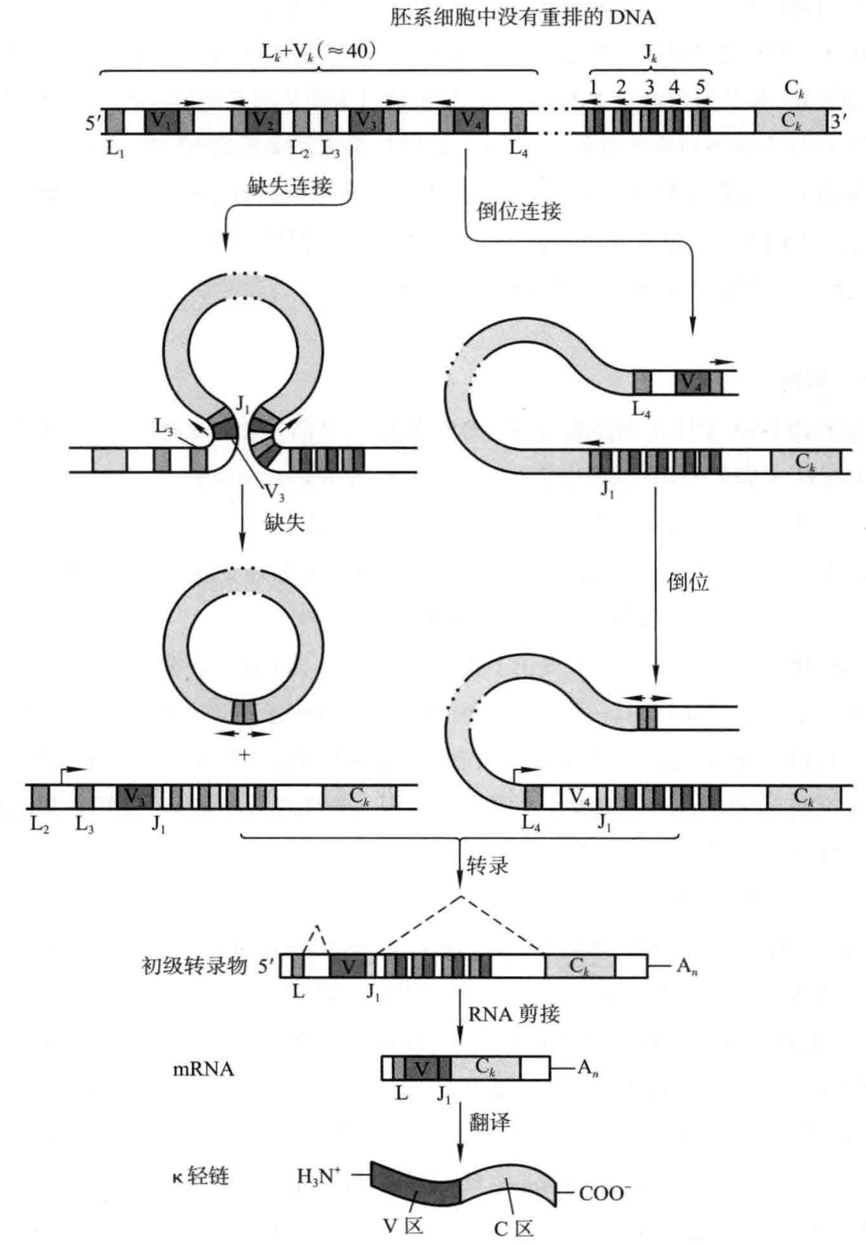


图 43-4　抗体轻链基因的重排机制

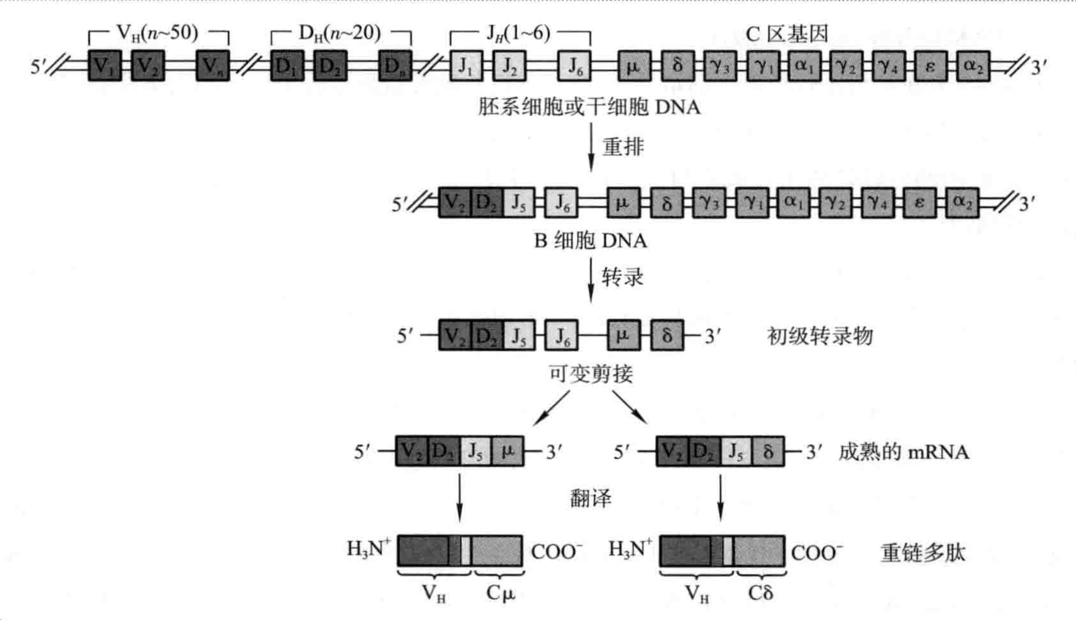


图 43-5　抗体重链基因的结构及其重排、转录、后加工和翻译

抗体基因的重排属于位点特异性 DNA 重组。通过重排,不仅产生了抗体的多样性,还将抗体基因的启动子(与 V 基因相邻)带到增强子(位于 J 基因和 C 基因之间)附近,从而使抗体基因能够有效地转录。

抗体基因的多样性除了通过 DNA 重排和 TdT 的作用创造以外,还可以通过 mRNA 前体的可变剪接、可变加尾、体细胞突变等手段来加强,概括起来如下:①VJ 和 VDJ 的重组连接。包括 κ 轻链的 40 个 V 基因和 5 个 J 基因的组合,共有 200 种可能;λ 轻链的 40 个 V 基因和 4 个 J 基因,共有 160 种可能;重链的 50 个 V 基因、20 个 D 基因和 6 个 J 基因,共有 6 000 种可能。轻链和重链的随机组合可形成含 κ 轻链和 λ 轻链的抗体分别是 $200 \times 6\,000=1.2 \times 10^6$ 种和 $160 \times 6\,000=0.96 \times 10^6$ 种,总共有 2.16×10^6 种可能性。②连接的多样性。重排过程中的连接是不精确的,这种“故意”的不精确连接可导致氨基酸的变化或缺失,从而影响到抗原结合位点的结构,实际上它也是抗体上出现高变区的原因。③插入的多样性。TdT 作用导致 DJ 连接或 V-DJ 连接中随机插入若干个核苷酸到 V 基因、D 基因或 J 基因的末端。④体细胞超突变(somatic hypermutation),主要发生在 V 基因。⑤可变剪接和可变加尾(参看“转录水平后调控”)。

Quiz4 试想出一种体细胞发生超突变的机制。

三、DNA 甲基化与羟甲基化

DNA 甲基化是一种复制后加工反应,是真核生物发生表观遗传的又一重要因素。真核生物和原核生物的 DNA 都可以进行甲基化修饰,但两者的甲基化位点和甲基化功能完全不同。

真核生物 DNA 的甲基化位点主要是 CpG 二核苷酸(少数是 CpNpG 三核苷酸序列)之中的 C,甲基供体为 SAM,由 DNA 甲基化酶催化。C 被甲基化后成为 5– 甲基胞嘧啶。

CpG 序列在基因组中的分布并不均一,通常成簇存在,形成所谓的 CpG 岛(CpG island)。每一个 CpG 岛长度在 1~2 kb 左右,内有多个 CG 序列,通常位于基因的启动子附近或内部,并有可能延伸到基因的第一个外显子。

甲基化样式与基因表达有关:活性基因的 CpG 岛处于去甲基化状态,非活性基因的 CpG 岛处于甲基化状态。管家基因的 CpG 岛在所有的细胞都呈去甲基化状态,而组织特异性基因的 CpG 岛只在表达它的细胞才处于去甲基化状态。

Quiz5 与其他二核苷酸序列相比,CG 序列更容易发生自发点突变,这是为什么?

甲基化导致基因转录活性的丧失有 3 种可能的机制(图 43–6):①甲基化直接阻止了对甲基化敏感的转录因子(如 Ap–2、E2F 和 NF-κB)、RNA 聚合酶与启动子的结合,而导致转录不能进行;②甲基化胞嘧啶结合蛋白(methyl cytosine binding protein,MeCP)与甲基化位点结合,致使转录无法进行。有 2 种 MeCP,一种是 MeCP–1,另一种是 MeCP–2。MeCP–1 与甲基化位点的结合可阻止转录因子、RNA 聚合酶与启动子的结合,从而使转录无法启动,而 MeCP–2 与甲基化位点的结合可将其他转录阻遏蛋白(如 SIN3 和 HDAC)招募到启动子周围,从而也能阻止基因的转录。③DNA 甲基化改变了染色质结构,致使甲基化周围的染色质成为无转录活性的异染色质。这可能是识别并结合甲基化位点的蛋白质,将 HDAC 和促进染色质浓缩的重塑因子招募到修饰位点附近所引起的。

雌性哺乳动物在胚胎发育的早期,即在原肠胚形成的时候,外胚层细胞中的一个 X 染色质经过浓缩和甲基化而随机地灭活。灭活至少与 2 种较长的 ncRNA 有关:一种是由 X 染色质上的一个叫 *XIST* 的基因编码,即 Xist RNA(约 17 kb),另外一种为 Tsix RNA(约 50 kb),它有部分序列与 Xist RNA 互补。具体的灭活过程大致如下:①在胚胎发育的早期,两个 X 染色质上的 *XIST* 基因都有转录活性,但转录出来的 Xist RNA 会很快水解。②其中有一条 X 染色质开始表达 Tsix RNA。受 Tsix RNA 和其他一些因素的作用,这一条 X 染色质在 *XIST* 基因的启动子附近的 CpG 岛因发生甲基化而不再表达,它将来会处于活性的状态。③另外一条 X 染色质上的 *XIST* 基因则继续转录。转录出来的 Xist RNA 的前体经过后加工以后(剪接和加尾),沿着表达它的 X 染色质积累,但并不扩散到另外一条 X 染色质上,而是包被在这条 X 染色质上,通过招募一些蛋白质复合物来阻断其他基因的表达,并逐步诱导此染色质

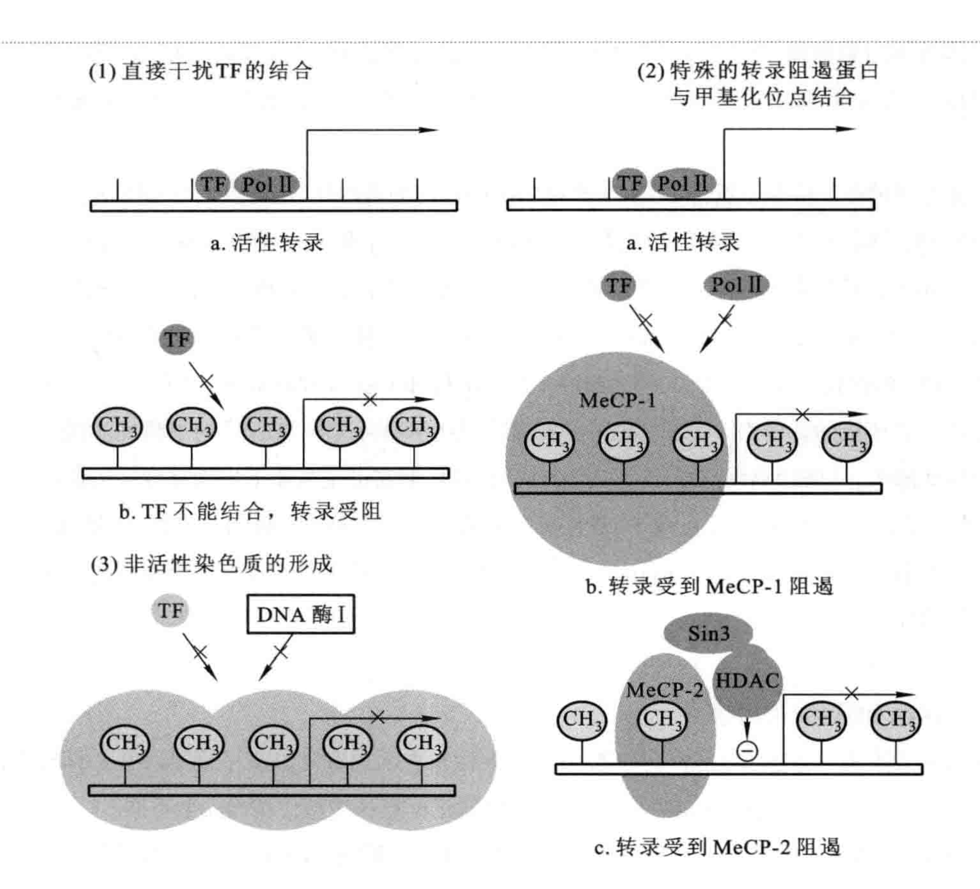


图 43-6 DNA 甲基化导致基因转录活性丧失的 3 种可能机制

Quiz6 如果一种猫的花色由 X 染色体上的基因控制，那么由这种猫的雌性和雄性克隆出来的猫与原来的猫在花色上是否相同？为什么？

浓缩为无转录活性的巴氏小体。受到招募的蛋白质包括组蛋白的一种变体——H2A1.2 以及 HDAC 和组蛋白甲基化酶。受 Xist RNA 招募的复合物在 X 染色体上只有一个进入点，就是 *XIST* 基因本身。④经过随机失活的细胞将来分裂的时候，形成的子细胞内的 X 染色质会维持原来的失活样式。

甲基化样式是可遗传的，它关系到某一物种一些特殊表型的产生（参看下文“DNA 印记”）。此外，甲基化样式具有组织特异性，这与组织特异性去甲基化酶有关系。

羟甲基化修饰也是发生在 DNA 的 C 上，C 因此变成 5- 羟甲基胞嘧啶（5-hmC）。这种修饰早在 1952 年就在某些噬菌体 DNA 上被发现，但直到 2009 年，人们才发现哺乳动物的脑细胞和胚胎干细胞中大量存在这种修饰的碱基。在哺乳动物细胞内，5-hmC 是由 5 甲基 C 氧化而成，催化此反应的酶属于 Tet 蛋白家族。C 羟基化修饰的功能也是调节基因的表达，特别是在哺乳动物的脑细胞和胚胎干细胞中，此外还能促进 C 的去甲基化，从而影响到 C 甲基化对基因表达的阻遏作用。

四、DNA 印记

就许多真核生物的某些基因而言，在一个发育的个体之中，两个等位基因中只有一个表达，有的是来自父本的基因才能表达，有的是来自母本的基因才能表达。这种由亲代决定的等位基因选择性表达的机制称为 DNA 印记（DNA imprinting）。印记的手段是甲基化，不表达的等位基因在 CpG 岛上的 C 被甲基化，表达的等位基因在 CpG 岛上的 C 没有发生甲基化。

Quiz7 你能否使用 DNA 印记来解释哺乳动物难以进行孤雌生殖？

在配子形成（gametogenesis）时期，许多基因就开始以性别特异性的方式进行甲基化反应，性别特异性的甲基化导致胚胎内来自不同亲本的等位基因的区别表达（图 43-7）。虽然在胚胎发育的早期，其生殖细胞内的甲基化样式重新设定，需要经历一波又一波的去甲基化和再甲基化反应，但这并不影响到印记基因的甲基化。

在个体发育的整个阶段，由于组织特异性甲基化酶的作用，不同类型细胞的甲基化样式会发生改变，但被印记的基因始终得到维持，这要归功于细胞内一种维持甲基化酶（maintenance methylase）的作用。

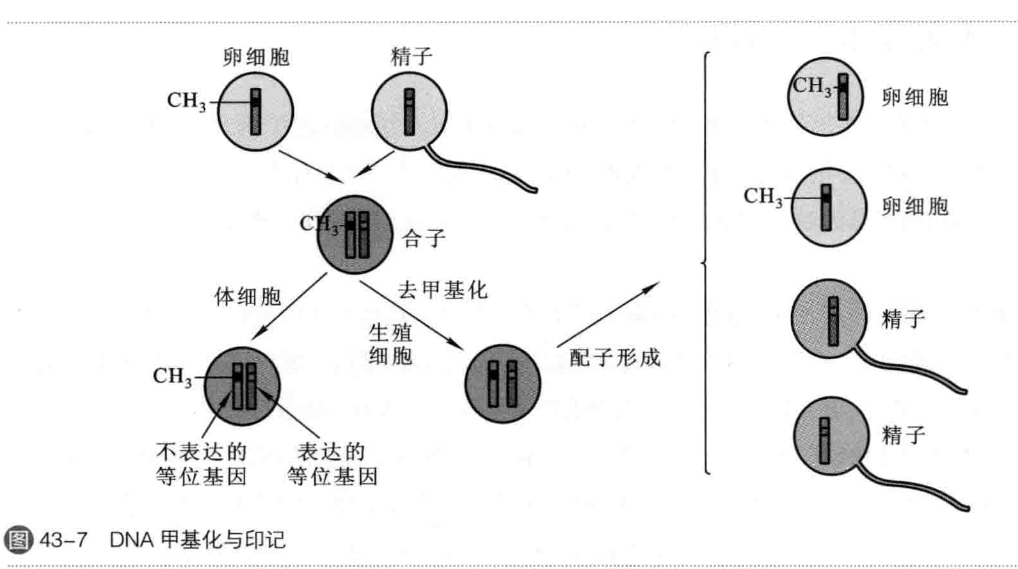


图 43-7 DNA 甲基化与印记

框 43-1 身边的生物化学——蜂王称王的秘密

大多数人都知道蜂王和工蜂的差别：蜂王有旺盛的生育能力，它唯一的工作是产卵，一天能下几千个卵，而工蜂却完全丧失了生育力；蜂王更长寿，能活几年，而工蜂只能活几个月。但蜂王是如何诞生的呢？

澳大利亚的 Ryszard Maleszka 的研究表明，蜂王是造出来的，不是天生的。当一群蜜蜂中需要一个新蜂王的时候，这时有一个幼虫被选中，只吃由保育工蜂头部腺体分泌出来的蜂王浆，它将来发育成蜂王，而其他幼虫只是在出生的时刻尝一口蜂王浆的味道以后，就改吃蜂蜜了，它们将来只能发育成工蜂。

为什么蜂王和工蜂有着相同的基因，结局却如此不同呢？根据 Ryszard Maleszka 等人发表在 2008 年 3 月 13 日 *Science* 上一篇论文的研究成果，原来这和“DNA 分子上 CpG 岛上的 C 甲基化”有关。DNA 的甲基化是由甲基化酶来催化，如 Dnmt3。如果让蜜蜂幼虫中这种酶失去作用，蜜蜂幼虫就发育成了蜂王，这与喂它蜂王浆的效果是一样的。可见蜂王浆的作用就是让控制 DNA 甲基化的酶不起作用。Maleszka 使用 Dnmt3 的 siRNA，让 Dnmt3 不能翻译出蛋白质。当 Dnmt3 的 siRNA 被注射到蜜蜂幼虫体内以后，能够模拟蜂王浆的作用，结果 72% Dnmt3 表达受阻的幼虫变成了蜂王的样子。它们具有发育完好的卵巢，其结构与正常通过蜂王浆发育而来的蜂王没有什么差别，而注射阴性对照 siRNA 的幼虫 77% 发育成工蜂。

再比较成年工蜂和蜂王基因组 DNA 在甲基化上的差别，结果发现基因组上，工蜂比蜂王约多了 600 个基因因为发生了甲基化而不能表达。

五、多个启动子的可变使用

真核生物的很多基因具有不止一个启动子，在转录的时刻可选用不同的启动子。启动子的可变使用一般具有不同发育阶段的特异性或组织特异性，即在不同的发育阶段或者在不同类型的组织细胞中使用不同的启动子，从而导致一个基因可以编码出多种不同的 mRNA。据估计，人和小鼠超过一半以上的基因具有选择性启动子。例如，在哺乳动物体内，编码翻译起始因子 eEF1A 的基因和编码血红蛋白 γA 的基因在胚胎发育期间以及胚胎发育之后，分别使用无 TATA 盒和含有 TATA 盒的启动子来驱动转录。再如，编码羟色胺 3 型受体 B 亚基的基因有 2 个不同的启动子，分别在外周神经系统和中枢神经系统中使用。

Quiz8 真核生物有哪些机制可以让一个基因编码出多种蛋白质或多肽？

同一个基因受不同启动子驱动而转录出来的 mRNA 可能具有不同的 5′UTR，也可能具有不同的 ORF，它们经过翻译可产生不同性质或功能的蛋白质产物。真核生物通过这种方式可以提高一个基因的编码能力，以增加蛋白质的多样性。

第三节 在转录水平上的调控

真核生物的蛋白质基因的转录除了启动子、RNA 聚合酶Ⅱ和基础转录因子以外，还需要其他顺式作用元件和反式作用因子的参与。参与基因表达调控的主要顺式作用元件有增强子、沉默子、绝缘子和各种应答元件；参与基因表达调控的反式作用因子有介导蛋白、激活蛋白、辅激活蛋白、阻遏蛋白和辅阻遏蛋白等。

激活蛋白与增强子结合，可以激活基因的表达，但需要辅激活蛋白才能起作用；阻遏蛋白与沉默子结合，抑制基因的表达，但需要辅阻遏蛋白才能起作用。某些转录因子既可以作为激活蛋白，也可以作为阻遏蛋白起作用，而究竟是起何种作用则取决于被调节的基因。辅激活蛋白缺乏 DNA 结合位点，但它们能够通过蛋白质与蛋白质的相互作用来行使功能，作用方式包括：招募其他激活蛋白、组蛋白修饰酶（如 HAT）和染色质重塑因子（如 SWI/SNF）到转录复合物而协助激活蛋白激活基因的转录；辅阻遏蛋白也缺乏 DNA 结合位点，但同样可通过蛋白质与蛋白质的相互作用起作用，作用机理包括掩盖激活蛋白的激活位点、作为负别构效应物和携带去修饰酶（如 HDAC）去中和修饰酶的活性。

一、顺式作用元件

1. 增强子

作为激活蛋白的结合位点，增强子在激活基因表达过程中起着不可替代的作用。核小体的存在使体积庞大的 TFIID 和 RNA 聚合酶Ⅱ很难与启动子结合，然而，激活蛋白具有简单而有效的 DNA 结合位点，通过这些结合位点与增强子的结合，可“扰乱”局部的染色质结构，从而暴露出一个基因的启动子或者它的调节位点，为基因表达铺平道路。

2. 沉默子

沉默子的结构特征和作用特点与增强子极为相似，只是作用效果正好与增强子相反。它作为阻遏蛋白的结合位点，可抑制特定基因的表达。

3. 应答元件

应答元件是细胞为了应对各种信号刺激做出反应而存在于 DNA 分子上特殊的核苷酸序列。这些特殊的序列存在于某些基因的上游，与增强子、沉默子和启动子一起调节其下游基因的表达。常见的应答元件有 HRE、HSE、MRE、SRE、CRE 等。

4. 绝缘子

绝缘子是一种位于启动子与增强子或启动子与沉默子之间的顺式作用元件，它具有 2 个重要功能（图 43-8）：①作为染色质上转录活性基因和非活性基因的界限，在与特殊的转录因子结合以后可防止激活或阻遏扩展到相邻的基因。②如果位于一个增强子（或沉默子）和一个启动子之间，绝缘子就可阻断增强子（或沉默子）对其下游基因转录的激活（抑制），但它并不干扰启动子的功能；如果它位于一个增强子（或沉默子）的上游，对这个增强子（或沉默子）的作用就无影响。

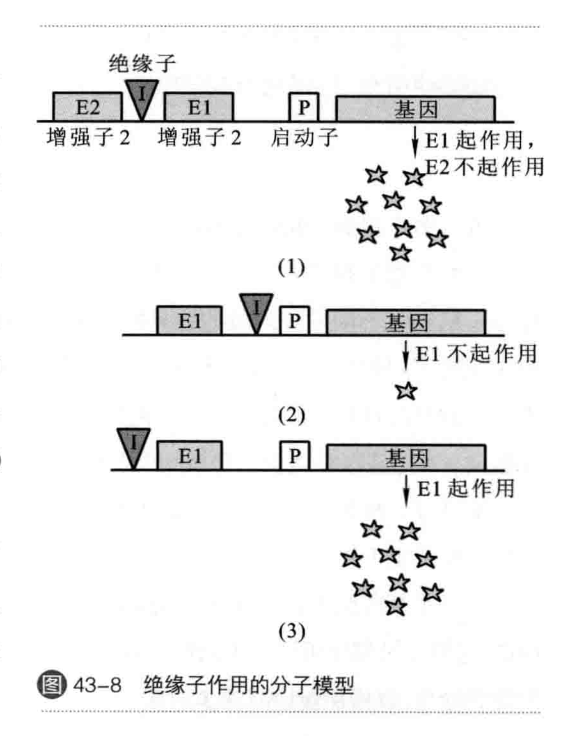


图 43-8 绝缘子作用的分子模型

绝缘子并不是不可改变的调控元件，其“绝缘”的效率常常受到增强子、沉默子、启动子和基因组上下游序列的影响。此外，绝缘子本身的功能也是受到调控的，并且还可能有组织特异性。

绝缘子还可能参与 DNA 印记：来自父本和母本的绝缘子一个甲基化，一个非甲基化，只有非甲基化的绝缘子才能与特定的转录因子结合，产生“绝缘”的效果。

以小鼠为例，其基因组上的印记控制区域（imprinting control region，ICR）就是一种绝缘子，它与类胰岛素生长因子（insulin-like growth factor 2，*Igf2*）的基因和 *H19* 基因的印记有关。*Igf2* 和 *H19* 相距约 90 kb，只有来自父本的 *Igf2* 和来自母本的 *H19* 才具有转录的活性。两个基因的转录可受到 *H19* 下游同一个增强子的激活。

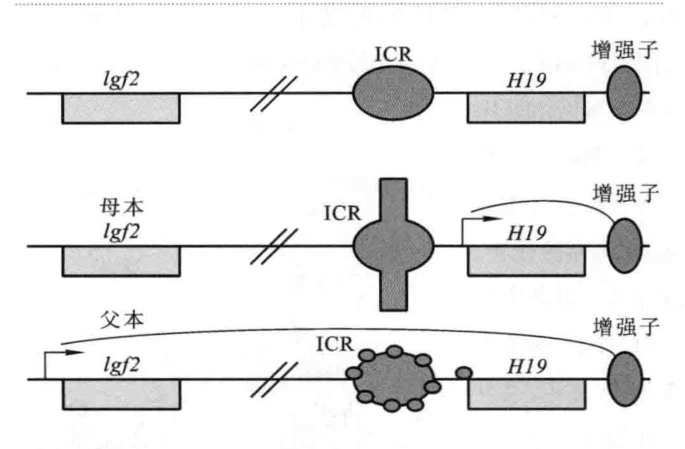


图 43-9　ICR 绝缘子对小鼠 *Igf2* 和 *H19* 基因表达的控制

如图 43-9 所示，ICR 位于 *Igf2* 和 *H19* 之间，靠近 *H19*，它富含 CG 序列，能够与 CCCTC 结合因子(CCCTC-binding factor，CTCF）结合，但只有非甲基化的 ICR 才能与 CTCF 结合。与 CTCF 结合的 ICR 才能够阻断增强子刺激绝缘子以外的基因转录。

来自父本的 ICR 及其周围的序列以及 *H19* 的启动子均处于甲基化状态，因而无转录活性，*H19* 无法转录，但 *Igf2* 却受到增强子的激活而具有转录的活性；来自母本的 ICR 及其周围的序列以及 *H19* 的启动子处于去甲基化状态，因而有活性，*H19* 受到增强子的激活而转录，但 *Igf2* 却因 ICR 的绝缘得不到增强子的激活而无转录活性。

二、转录因子

转录因子泛指除 RNA 聚合酶以外的一系列参与 DNA 转录和转录调节的蛋白质因子，分为基础转录因子和调节性转录因子。基础转录因子专指从核心启动子开始进行精确转录所必需的一组最低数量的蛋白质的总称，而其他转录因子参与转录的调节，激活或阻遏基因的转录，因此属于调控性转录因子。

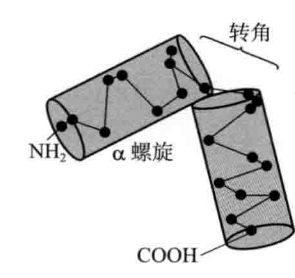
（一）转录因子的结构

到目前为止，已在不同的真核生物中发现近 2 800 多种转录因子，约占哺乳动物基因总数的 10%。大多数转录因子通常含有以下几种不同的结构域：①DNA 结合结构域（DNA-binding domain，DB），直接与顺式作用元件结合的转录因子都具有此结构域。转录因子一般使用此结构域之中的特殊 α 螺旋与顺式作用元件内的大沟接触，通过螺旋上特殊氨基酸残基的侧链基团与大沟中特殊碱基对形成的次级健（主要是氢键），来进行相互识别而产生特异性。许多转录因子在此结构域上富含碱性氨基酸，这可能有利于和 DNA 骨架上带负电荷的磷酸基团发生作用。②效应器结构域（effector domain），这是转录因子调节转录效率（激活或阻遏）、产生效应的结构域。③多聚化结构域（multimerization domain），此结构域的存在使得转录因子之间能够组装成二聚体或多聚体（同源或异源）。④还有的转录因子含有第 4 种即信号感应结构域(signal sensing domain)，通过这种结构域，转录因子能够探测到外部特殊的信号，并将这种信号传递给转录复合物，从而改变特定基因的表达。例如，固醇类激素的受体就是具有这种结构域的转录因子。下面将集中介绍前 2 种结构域，特别是 DNA 结合结构域。

1. DNA 结合结构域

DNA 结合结构域一般会含有以下几种结构模体中的一种（42-10）：

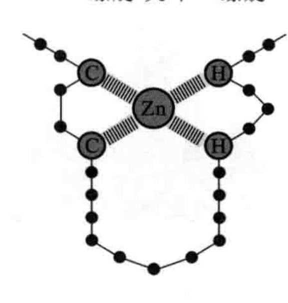
（1）α 螺旋 - 转角 -α 螺旋（helix-turn-helix，HTH）　这是第一个得到确定的 DNA 结合模体。它最初被发现在 λ 噬菌体的 Cro 蛋白，后来又在大肠杆菌的乳糖操纵子、色氨酸操纵子的阻遏蛋白以及 CAP 激活蛋白中发现。参与真核生物发育的被称为同源异形域蛋白质（homeodomain protein）的转录



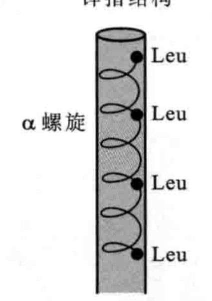
α 螺旋 - 转角 - α 螺旋



α 螺旋-突环-α 螺旋



锌指结构



亮氨酸拉链 α 螺旋

图 43-10　转录因子 4 种 DNA 结合模体

因子也有这种模体。

HTH 的主要特征包括：①由约 20 个氨基酸残基组成，这些氨基酸残基形成 2 个短的 α 螺旋（长 7~9 个氨基酸）和 1 个位于螺旋之间的由 4 个氨基酸残基组成的转角（非 β 转角）结构，转角通常起始于 Gly 残基；②两个 α 螺旋近乎垂直，其中的一个被称为 DNA 识别螺旋（DNA recognition helix）；③识别螺旋直接与 DNA 双螺旋的大沟接触，以识别和结合特异性的顺式作用元件，另外一个螺旋有助于识别螺旋在空间上采取合适的取向，并可能与转录因子与 DNA 结合的总的亲和性和稳定性有关系。

（2）碱性 α 螺旋 – 环 –α 螺旋（basic helix-loop-helix，bHLH） 该模体由 2 段 α 螺旋和螺旋之间的突环构成。两段 α 螺旋一小一大。小的具有两亲的性质，有利于两条肽链之间形成二聚体，大的位于 N 端，一般含有一簇碱性的氨基酸残基，由它负责识别并结合特殊的 DNA 序列。被结合的碱基序列称为 E 盒，其一致序列为具有回文性质的 CACGTG。两条多肽链可通过两亲的小 α 螺旋在疏水面之间的相互作用，来组装成同源二聚体或异源二聚体蛋白，每个单体则再通过大的 α 螺旋与 E 盒结合（图 43–11）。

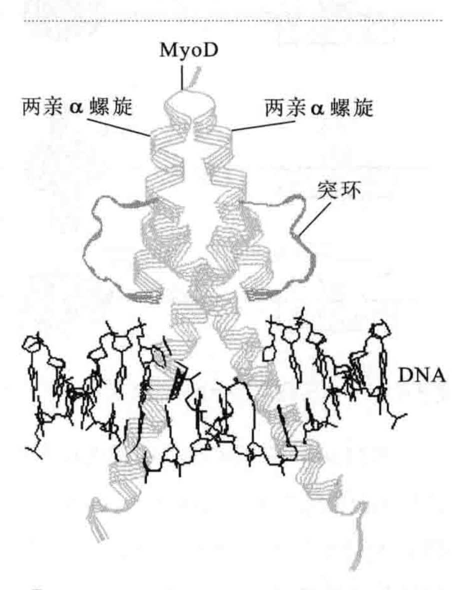


图 43–11 含有碱性 α 螺旋 – 突环 –α 螺旋结构域的 MyoD 与 DNA 的结合

许多 bHLH 类转录因子参与真核生物的发育，通常被分为 2 类：一类广泛表达，无组织特异性，如哺乳动物的 E12/E47；另一类具有组织特异性，如 MyoD 和 Myf5 仅在肌细胞中表达。

（3）锌指结构（Zinc finger） 该模体最初是在真核生物 RNA 聚合酶Ⅲ的基础转录因子 TFIIIA 中发现的。据估计，真核生物中含有这种结构的转录因子最多，约占哺乳动物基因总数的 1%。

锌指结构的主要特征是：①含有 Zn^{2+}；② Zn^{2+} 与肽链上若干个含有孤对电子的氨基酸残基的侧链（His 或 Cys）有规律的配位结合，使多肽链上一段相对短的序列以 Zn^{2+} 为中心折叠成一种致密的指状结构；③ Zn^{2+} 并不与 DNA 结合，而是指状结构中的 α 螺旋伸入到 DNA 双螺旋的大沟内，以识别和结合特异性的顺式作用元件。

需要特别指出的是，并不是含有锌指结构的蛋白质就一定能够与 DNA 结合，如含有锌指结构的 PKC 就不能与 DNA 结合。此外，某些含有锌指结构的蛋白质还能与 RNA 结合。

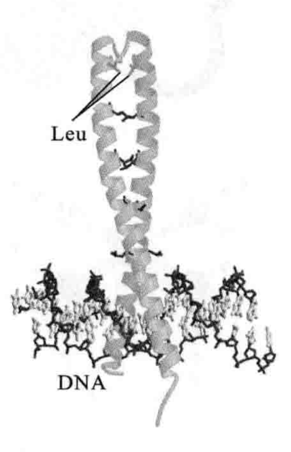


图 43–12 碱性拉链结构域与 DNA 的结合

（4）碱性拉链（basic zipper，B-Zip） 该模体由 1 段碱性的 DNA 结合肽段和相邻的亮氨酸拉链结构组成，其中亮氨酸拉链（leucine zipper）由 2 条多肽链上的 2 段两亲 α 螺旋组成，两段螺旋通过集中在一侧的 Leu 残基疏水 R 基团之间（每隔 6 个氨基酸残基出现 1 个 Leu 残基）的相互作用而结合在一起。

亮氨酸拉链的功能在于能使转录因子以二聚体的形式起作用。在形成二聚体后，两条肽链上带正电荷的碱性 DNA 结合肽段能够采取合适的取向，同时与 DNA 结合（图 43–12）。

具有 B-Zip 模体的蛋白质主要存在于真核生物中，其中三个最典型的转录因子是 AP–1、ATF-CREB 家族和 GCN4。其中，AP–1 是由 Jun 和 Fos 两种不同的蛋白质组成的异源二聚体；GCN4 为同源二聚体，参与调节酵母细胞内与氨基酸合成有关的基因表达；ATF-CREB 家族由 13 个不同的单体组成，可形成 78 种不同的异源二聚体。



图 43–13 翼式螺旋

（5）翼式螺旋（winged helix） 这是一种由 α 螺旋和 β 折叠组成的模体结构，其中 α 螺旋有 3 段，位于 N 端，β 折叠是由 3 个反平行的 β 股组成。β 折叠结构使得 α 螺旋呈翼状（图 43–13）。这种模体结构最初发现于哺乳动物肝细胞的转录因子 HNF–3γ，现在在其他与 HNF–3γ 同属一族的转录因子中也有发现。含有这种模体结构的 DNA 结合蛋白依赖于其中的 2 个 α 螺旋与 DNA 大沟中的序列相互作用。

(6) 与小沟接触的 β 支架因子(beta-scaffold factor with minor groove contact) 大多数转录因子与DNA 双螺旋的大沟接触,但含有这种模体的蛋白质却在小沟与 DNA 结合,识别碱基序列的不是 α 螺旋,而是突出在外的 β 折叠。具有这种模体的蛋白质有 STAT、TBP、p53 和 HMG 等。

2. 激活蛋白的效应器结构域

激活蛋白的效应器结构域即是激活基因转录的激活结构域(activation domain,AD)。转录因子上的 DNA 结合结构域能让转录因子与特定的顺式作用元件结合,以"锁定"被调节的目标基因,激活基因表达的功能由转录因子上专门的激活结构域承担。

转录因子通过其激活结构域与转录机器上特定的组分或者其他调节蛋白结合,从而影响转录的效率。

Quiz9 有时一种激活蛋白过量表达能导致其激活功能的丧失,甚至还能抑制细胞内其他激活蛋白的作用,为什么?

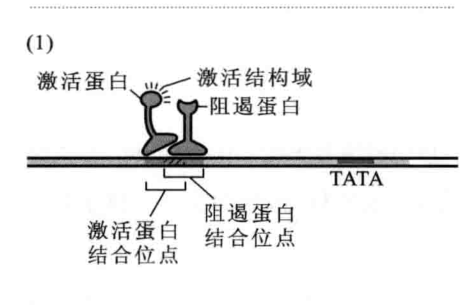
已发现 3 种常见的激活结构域:①酸性结构域(acidic domain),富含酸性氨基酸残基(Asp 和 Glu),但常有 1 个疏水的氨基酸残基镶嵌在其中。例如,糖皮质激素受体、酵母细胞的 GAL4 和 GCN4 蛋白。氨基酸替换实验显示,增加酸性结构域的负电荷数,能提高激活效率。②富含 Gln 结构域(glutamine-rich domain),如 Sp1、Oct-1、Oct-2、Jun 和 GAL11。③富含 Pro 结构域(proline-rich domain),例如 CTF/NF-1 家族和 AP2。

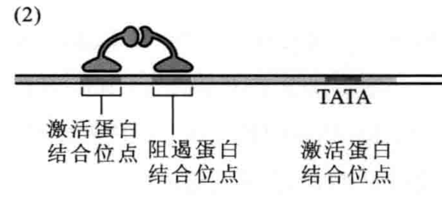
(二) 转录因子的分类

根据 DNA 结合结构域的特征,转录因子可分为 5 大类:第一大类为碱性结构域类转录因子,含有 bHLH、B-Zip 或 bHLH-Zip 模体的转录因子以及 NF-2、RF-X 和 b-HSH 类转录因子都属于此大类;第二大类为锌配位 DNA 结合结构域(Zinc-coordinating DNA-binding domain)类转录因子,含有锌指结构的转录因子均属于此类;第三大类为 HTH 类转录因子,含有 HTH 模体的转录因子均属于此类;第四大类为与小沟接触的 β 支架因子;第五大类称为零大类,不属于上述四大类的转录因子原则上可被归入此类,如 AP2/EREBP 相关因子和 HMGI。

(三) 转录因子本身的激活机制

大多数转录因子本身的活性也是受到调节的,这些转录因子在调节其他基因转录之前自己首先要被激活,激活的途径以下几种:①转录激活。某些转录因子本身的基因通常处于关闭的状态,只有在特定的条件下才被诱导表达,如转录因子 Fos 的表达受到磷酸化的血清应答因子(SRF)的诱导才会表达。②磷酸化。许多转录因子只有在受到蛋白质激酶的作用才有活性,如 HSF 和 CREB 等。③去磷酸化。某些转录因子只有在磷酸酶作用下,从磷酸化状态转变为去磷酸化状态才有活性。④配体结合。脂溶性激素的受体只有在与脂溶性激素结合(充当配体)以后,才能够作为转录因子调节某些基因的表达。⑤对象交换(partner exchange)。一个二聚体转录因子可能有不同的聚合对象,与不同的聚合对象组装成的二聚体活性会有变化,其中某些二聚体有活性,某些二聚体没有活性,无活性的二聚体只有在其聚合的对象被交换以后才被激活。例如,转录因子 Myc、MAD 和 MAX,MAX-MAX 同源二聚体的活性很低,MAD-MAX 能够与相应的顺式作用元件结合,但却不能激活基因的转录,只有 Myc-MAX 才有活性。⑥抑制剂释放。某些转录因子本来与细胞内内源的抑制剂结合而无活性,只有当它与抑制剂解离以后才有活性。





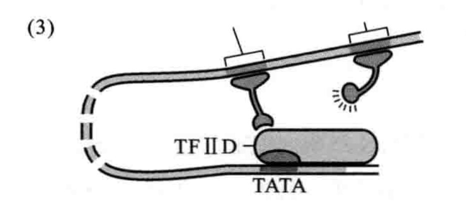


图 43-14 阻遏蛋白与沉默子结合以后抑制基因表达的 3 种可能方式(Alberts 等,2007)

(四) 阻遏蛋白的抑制机制

阻遏蛋白与沉默子结合以后可以通过 3 种方式来抑制基因的表达(图 43-14):①与激活蛋白竞争结合增强子的部分序列,这是因为增强子有部分序列与沉默子重叠,因此,

Quiz10 若是将一种激活蛋白的激活结构域与一种阻遏蛋白的DNA结合结构域结合，那么这种融合蛋白在与相应的沉默子结合以后，究竟是激活还是抑制相关基因的表达？

阻遏蛋白与沉默子的结合将会干扰激活蛋白与增强子的结合；②与激活蛋白的激活结构域结合，使激活蛋白不能起作用；③直接作用于基础转录因子或RNA聚合酶。

与激活蛋白相似，许多阻遏蛋白也具有2个结构域，一个是DNA结合结构域，另一个则是阻遏结构域（repression domain）。

三、转录水平调控的实例

（一）酵母细胞半乳糖代谢相关基因的表达调控

酵母细胞内参与半乳糖代谢的3个基因 *Gal1*（半乳糖激酶基因）、*Gal7*（半乳糖转移酶基因）和 *Gal10*（半乳糖差向异构酶基因）的表达受到半乳糖的诱导，这3个基因虽然靠得很近，但并不像原核生物那样组成操纵子。

在 *Gal1* 和 *Gal10* 之间有一段上游激活子序列（upstream activator sequence，UAS），其一致序列是 CGG-N_{11}-CCG（N为任何一种碱基），为转录因子GAL4蛋白的结合位点。

在没有半乳糖时，GAL4蛋白二聚体与GAL80蛋白组成的复合物与UAS结合，这时GAL80与GAL4的激活结构域结合，而阻止GAL4去激活 *Gal* 基因的转录；在有半乳糖时，半乳糖的代谢物与GAL80上的结合位点结合，诱使GAL80与GAL4解离并与细胞质中的Gal3结合，GAL4的激活结构域因此暴露出来，于是，它能将一系列促进和参与转录的蛋白质，包括SAGA（具有HAT活性）、介导蛋白、基础转录因子和RNA聚合酶Ⅱ招募到 *Gal* 基因的启动子周围，激活 *Gal* 基因的表达（图43-15）。

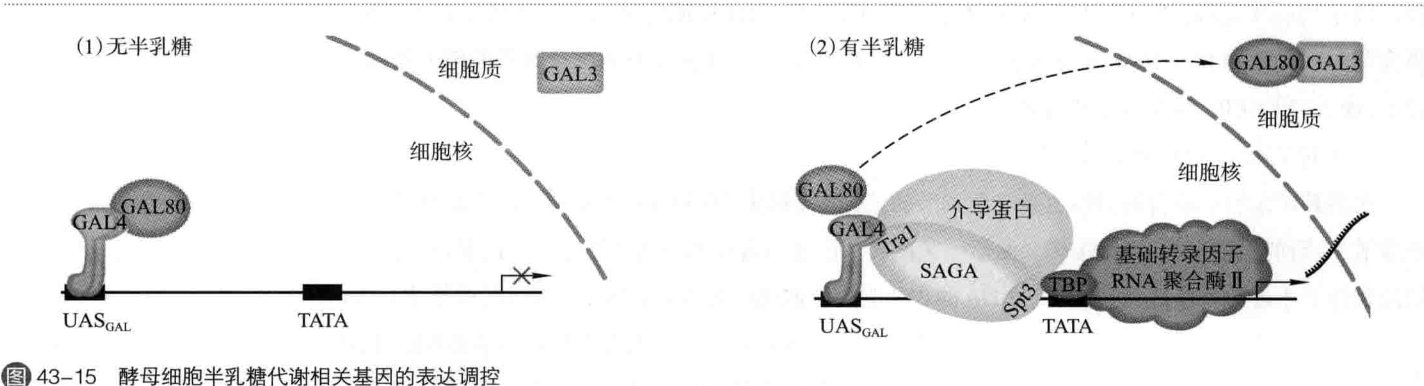


图43-15 酵母细胞半乳糖代谢相关基因的表达调控

（二）热激蛋白的基因表达调控

真核生物在温度骤然升高或其他应激状态下，体内HSP的基因被诱导表达。HSP的基因表达除了需要启动子以外，还受到HSE和热激因子（heat shock factor，HSF）的控制，其中HSE位于热激基因的上游，HSF能识别HSE内特异性的-NGAAN-序列。

就哺乳动物而言，无应激时，HSF以单体形式存在，不能结合DNA；而在应激状态下，HSF从单体变成三聚体，随后进入细胞核，并与HSE结合，通过它的激活结构域将p100/CBP复合物（具有HAT活性）、介导蛋白、基础转录因子、RNA聚合酶Ⅱ招募到HSP的启动子周围，激活HSP的表达。HSF形成三聚体的时候，相互之间通过寡聚体结构域结合，具体位点是寡聚体结构域内的3段七肽重复序列。无应激时，这3段七肽序列在HSF单体内部形成稳定的卷曲螺旋结构；应激时，卷曲螺旋结构打开，相邻HSF的七肽重复序列相互作用，形成分子间的卷曲螺旋，于是，3个HSF单体结合在一起形成一个三聚体。

（三）金属硫蛋白的基因表达调控

金属硫蛋白（metallothionein，MT）是一种富含Cys残基的小蛋白，在细胞内能够与重金属离子（如铬、锌、铜和汞等）结合，帮助细胞清除过量的重金属，从而保护细胞免受重金属毒害。此外，它还参与细胞防护活性氧以及调节体内锌离子浓度的稳定。

MT基因平常以较低的水平表达，但在过量的重金属离子或在糖皮质激素的作用下，可大量表达。

重金属离子可诱导 MT 的表达是因为 MT 的上游存在 MRE(一致序列是 TGCRCNC),而 MRE 需要金属应答性转录因子 -1(metal-responsive transcription factor 1,MTF-1)的结合。

MTF-1 是高等真核生物体内唯一一种可探测重金属水平的转录因子,而锌是唯一能够可逆结合并激活 MTF-1 的金属离子。MTF-1 的 N 端含有 DNA 结合结构域,具有 C_2H_2 锌指结构,而它的 C 端同时含有 3 种转录激活结构域,即酸性结构域、富含 Pro 结构域和富含 S/T 的结构域。

当细胞内的锌离子过量的时候,便与 MTF-1 结合,MTF-1 因此被激活。而一旦 MTF-1 被激活,即与 MRE 结合,然后通过它的激活结构域将 p100/CBP 复合物(具有 HAT 活性)、介导蛋白、基础转录因子、RNA 聚合酶Ⅱ招募到 MT 的启动子周围,激活 MT 的表达。

(四) 生物发育过程中的组织特异性基因表达

组织特异性基因表达与组织特异性转录因子以及组织特异性顺式作用元件有关,其主要特征反映在以下几个方面:①组织特异性转录因子只在特定类型的细胞内表达,通常呈时空特异性的表达;②组织特异性的转录因子与组织特异性的顺式作用元件结合,与基础转录因子一起调节基因的表达;③某些转录因子专门在细胞分化的早期阶段表达,作用于受阻遏的染色质上的启动子,并招募其他蛋白质,促进染色质的重塑,以形成具有转录活性的染色质结构;④转录激活通常还受到其他信号的调节;⑤在整个个体发育阶段,一种转录因子的表达往往受到另一种转录因子的调节,不同的转录因子会构成一种复杂的级联网络。

以果蝇的个体发育为例,在它的受精卵形成以后,需要依次经过卵裂(cleavage)、原肠胚(gastrula)、器官发生(organogenesis)和成熟与生长 4 个阶段。而在每一个阶段,都需要启动特定的基因表达程序。这种程序是高度可重复的,即对每一个胚胎都是相同的。这样才能保证由每一个受精卵发育成的个体都会在正确的位置分化出大小、形状和功能完全正确的器官。

e43-2 果蝇发育过程中的组织特异性基因表达

(五) 转录终止阶段的调控

真核细胞 DNA 转录的终止机制尚不清楚,但已发现某些基因的转录会发生提前终止(如 *c-myc* 基因)或抗终止现象(HIV 的 Tat 蛋白基因)。

第四节 在转录后加工及运输水平上的调控

转录后加工水平上的基因表达调控对于真核生物具有非同寻常的意义,它是一个基因产生多种多肽或蛋白质产物的主要机制。

一、可变剪接

可变剪接是指一种 mRNA 前体在剪接反应中某些区段的序列可能被保留,也可能被排除,从而得到几种不同成熟 mRNA 产物的过程。这是高等真核生物蛋白质多样性产生的重要来源。据估计,人类基因组中的基因至少有 60% 经历可变剪接,平均每个基因有 3~4 个剪接变体。

可变剪接主要有 4 种方式(图 43-16):①外显子跳过(exon skipping)——剪接反应中跳过一个或几个外显子,从而导致成熟的 mRNA 上缺失相应的外显子;②内含子保留(intron retention)——一个或几个内含子被保留下来而出现在成熟的 mRNA 之中;③可变的 3′- 剪接位点的使用——3′- 剪接位点不止一个,使用不同的剪接位点可产生不同的剪接产物;④可变的 5′- 剪接位点的使用——5′- 剪接位点不止一个,使用不同的剪接位点可产生不同的剪接产物。

可变剪接反应中剪接位点的选择受到多种反式作用因子和位于 mRNA 前体上特殊的顺式作用元件的调控。顺式作用元件通常是一段较短的核苷酸序列(8~10 nt),根据对剪接反应的不同影响,可分为剪接增强子(增强对某个剪接位点的使用)和剪接沉默子(抑制对某个剪接位点的使用)。剪接增强

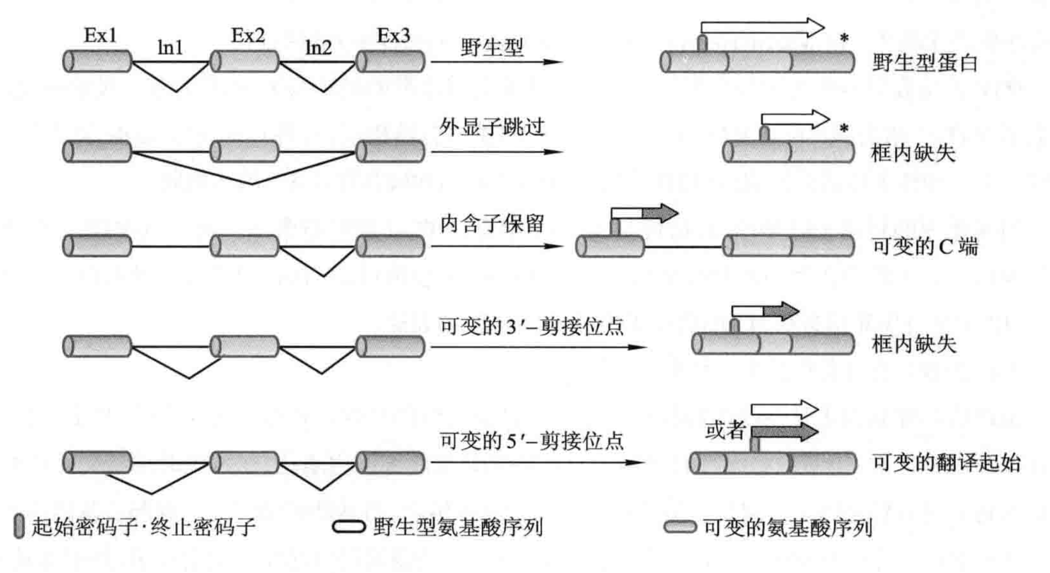


图 43-16　可变剪接的四种方式(Ex 和 In 分别表示外显子和内含子)

子和沉默子可能位于外显子或内含子之中,反式作用因子通过与顺式作用元件的结合而参与对不同剪接位点的选择。

调控可变剪接的反式作用因子有 2 类:基础剪接因子和特异性剪接因子。基础剪接因子间的协同作用和颉颃作用以及相对浓度的改变可以影响到剪接位点的选择,特异性剪接因子可以调控特异的剪接过程。剪接增强子多位于它们所调节的剪接位点附近,有助于吸引剪接因子到剪接位点上,改变剪接增强子的位置可使剪接活性发生很大改变,甚至使它们转变成为负调控元件。一类富含嘌呤核苷酸的外显子剪接增强子是最常见的一种剪接增强子。

可变剪接导致一个基因编码出不同的 mRNA。那些影响到 mRNA 编码氨基酸序列的可变剪接将会产生序列不同和活性不同的蛋白质变体,而发生在 mRNA 非编码区域的可变剪接可能会影响到成熟的 mRNA 稳定性或翻译的效率,具体表现在以下 5 个方面:①产生细胞定位不同的蛋白质。如神经细胞黏附分子(neural cell adhesion molecule,NCAM)的胞内型和胞间型、抗体和退化加速因子(decay-accelerating factor,DAF)的分泌型和膜结合型、纤黏蛋白(fibronectin)的胞内型和血浆型都是与可变剪接有关。②改变可读框,导致蛋白质活性的缺失。例如,果蝇性别决定有关的 Sxl(sex-lethal)的有活性型和无活性型的产生与此有关。③改变蛋白质的活性。有的是细调某些蛋白质的活性,如原肌球蛋白或蛋白质激酶和具有不同电生理活性的离子通道。④产生具有全新活性的蛋白质。例如,降钙素和降钙素基因相关肽的形成。⑤影响 mRNA 的稳定性和翻译的效率。如集落刺激因子 -1 mRNA 的 3′UTR 去稳定元件是否保留与可变剪接有关。

二、可变加尾

很多真核生物基因的 3′端含有 2 个或 2 个以上的加尾信号,使用不同的加尾信号将产生不同长度或不同性质的 mRNA。同一种 mRNA 前体在加尾反应中,对不同加尾信号的选择而产生不一样的成熟 mRNA 的现象称为可变加尾。

可变加尾可能会改变编码区的长度,也可能会保留或去除位于 3′UTR 内影响 mRNA 稳定性的特定信号。前一种情形是导致一个基因编码不同的多肽产物的另外一种途径,有时它与可变剪接组合使用可进一步扩大蛋白质的多样性。生物信息学研究表明,高等生物约 25% 的基因可能发生可变加尾。

以抗体的 μ 型重链为例,它有分泌型和膜结合型(图 43-17),其 mRNA 前体共有 4 个加尾信号。分泌型使用的是位于最前端的加尾信号,而膜结合型使用的是第二个加尾信号。使用第二个加尾信号的结果是保留了位于 3′端的 M1 和 M2 两个外显子,而 M1 和 M2 编码的一段氨基酸序列为跨膜的 α 螺旋。

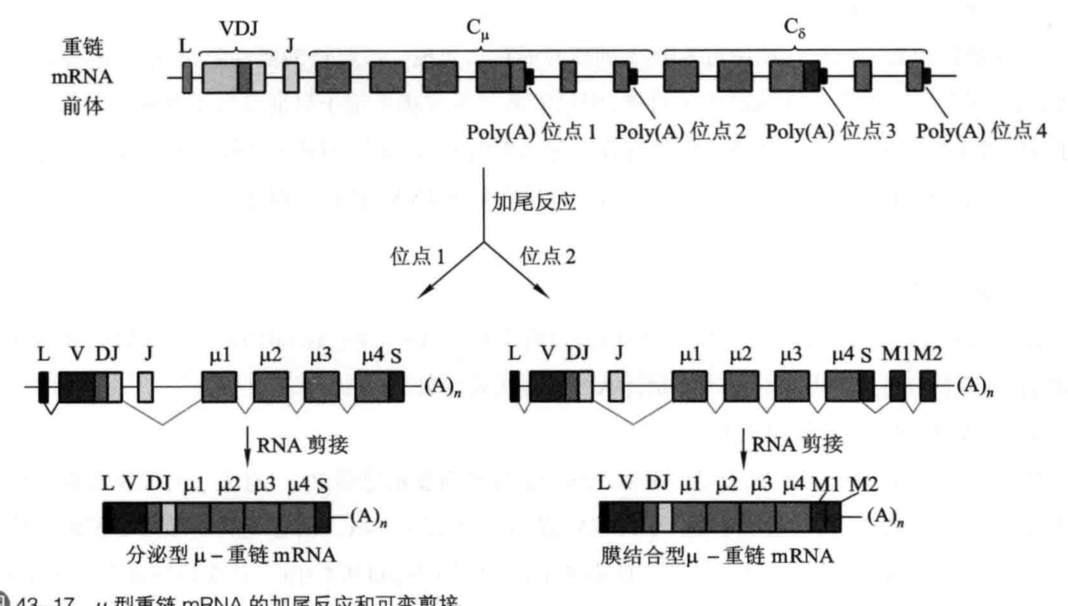


图 43-17 μ 型重链 mRNA 的加尾反应和可变剪接

三、在 mRNA 运输水平上的调控

真核生物基因的转录和翻译分别发生在细胞核和细胞质。mRNA 先在核内进行转录和后加工，然后再运输到细胞质作为翻译的模板。RNA 从细胞核运输到细胞质的过程受到严格的调控。例如，只有剪接好的 RNA 分子，而不是 mRNA 前体或还保留套索结构的 RNA，才会被运输出细胞核。显然，这种调控可防止异常的或具有潜在毒性的蛋白质被合成。

近年来，已分别在酵母和人细胞内发现将 RNA 剪接和 RNA 运输偶联在一起的蛋白质。这些蛋白质的功能是与剪接过的 RNA 结合，并将它们引向核孔。

有关组织特异性 RNA 编辑的内容请参看第三十七章“转录后加工”。

第五节　在翻译及翻译后加工水平上的调控

一、自体调控

自体调控通常被用来控制大分子聚合体的合成，它是一种调节大分子复合体内某一成分的合成以实现各成分能够协调表达的常见手段。以构成微管的前体——微管蛋白（tubulin）为例，自由的微管蛋白能够抑制自身 mRNA 的翻译。当微管装配受阻的时候，游离的微管蛋白亚基积累，并可能与自身的 mRNA 结合，或与新生的肽段结合，从而导致位于多聚核糖体上的微管 mRNA 发生降解，以抑制多余微管蛋白的合成。

二、mRNA 区域化

在很多细胞内，特别是卵母细胞，mRNA 被限制在特定的区域，一个典型的例子已在果蝇的胚胎发育中讨论过。其他的例子包括：正在迁移的成纤维细胞内的肌动蛋白 mRNA 被限定在细胞的前端；神经元内的 Tau mRNA 定位于轴突，MAP2 mRNA 定位于胞体及树突，因而 Tau 的表达限制在轴突内，而 MAP2 只在胞体和树突表达。由于 Tau 和 MAP2 在细胞内的功能是决定微管的行为，因此这两种蛋白质 mRNA 的选择性分布控制了细胞骨架在神经元内的布局。

三、mRNA 的屏蔽

某些基因的表达调控可通过对 mRNA 的暂时屏蔽而实现。以雌性两栖动物为例，它的任意一个成熟的卵细胞已基本停止新 mRNA 的合成，但已贮备了所有将来用于早期发育的各种 mRNA。这些提前合成好的 mRNA 与特殊的蛋白质结合在一起而暂时被"屏蔽"，因此无法作为模板进行翻译。然而，一旦卵细胞受精并分裂，即进行解屏蔽反应，使贮备的 mRNA 可以进行翻译。

四、RNA 干扰

RNA 干扰（RNA interference，RNAi）是指通过特定双链 RNA 使目标 mRNA 发生降解或者翻译受到阻遏，从而特异性地抑制目标基因在翻译水平上表达的现象。

1. RNAi 的发现及其两种形式

RNAi 的发现可算得上是近十多年来生命科学领域最重要的进展之一（有关它的发现历程详见本章科学故事）。事实上，有 2 种类型的干扰 RNA，即 miRNA 和 siRNA。两者成熟的形式都是短的双链 RNA，每条链的长度通常为 21~25 nt。它们的来源不同，但作用机制基本相同，许多作用成分是共享的，最后的作用效果也是一致的。其中 siRNA 为外源的，可能来自病毒，也可能是人为导入产生的，其天然前体本来就是双链 RNA；而 miRNA 则由内源基因编码，其前体是内部带有发夹结构的单链 RNA，绝大多数由 RNA 聚合酶Ⅱ催化转录，只有少数由 RNA 聚合酶Ⅲ催化转录。

Quiz11 如何确定一种 miRNA 究竟由哪一种 RNA 聚合酶催化转录？

2. miRNA 的分布和基因结构及组织

miRNA 存在于绝大多数真核生物。据估计，人类基因组约编码大于 1 000 种的 miRNA，而受 miRNA 作用的基因约占蛋白质基因总数的 60%。某些病毒也编码 miRNA，例如，疱疹病毒约编码大于 140 种的 miRNA，这些 miRNA 约有 2/3 是控制自身的基因表达，其他控制宿主细胞的基因表达，特别是与宿主免疫有关的基因，此外还包括宿主细胞内的某些 miRNA。

在基因组上，miRNA 可成簇分布或独立存在，位于其他基因之间，或者在其他基因的反义区，也可能单独寄居在某种蛋白质基因或非蛋白质基因的内含子内，甚至在非蛋白质基因的外显子中。某些 miRNAs 在后加工的时候，会发生位点特异性的编辑。RNA 的编辑可阻止 pri-mRNA 的进一步加工，或者改变下游的加工方式从而增加 miRNA 的多样性。

3. miRNA 的命名

miRNA 的命名一般用 mir-X 或 miR-X 来表示，前者表示的是一种 miRNA 的前体，后者表示的是成熟的 miRNA，X 为数字编号，其大小通常表示发现的先后次序。有时还在一种 miRNA 前面加上代表某物种的前缀，如 hsa 代表人，因此 hsa-miR-194 代表的是来源于人细胞的一种 miRNA。一个成熟的 miRNA 由 2 条链组成（miRNA:miRNA*），但一般只有一条链去作用目标 mRNA，这一条链称为引导链（guide strand），另一条链不起作用，在细胞内容易水解，量较低，就像"过客"一样，故被称为过客链（passenger strand）。过客链可用带有星号的 miRNA 即 miRNA * 表示。

4. miRNA 的转录、加工和成熟

转录出来的首先是 miRNA 前体，即 pri-miRNA。Pri-miRNA 一般有几千个核苷酸，内部含有局部的发夹结构，若是后生动物（metazoan），就需要先后经过核加工、核输出、细胞质再加工，然后参入到 RNA 诱导的沉默复合体（RNA-induced silencing complex，RISC）之中，最后通过 RISC 锁定特定的目标 mRNA，并抑制其在翻译水平上的基因表达（图 43-18）。

Quiz12 若是果蝇体内的 Drosha 和 Pasha 蛋白缺失，那会影响 siRNA 的作用吗？

（1）pri-miRNA 的核加工　pri-miRNA 是一个大的前体，它首先在细胞核被加工成 70~80 nt 小的前体。这个小的前体称为 pre-miRNA，其二级结构实际上是一个大的发夹结构。催化这一步后加工反应的是一种约为 5.0×10^5 的蛋白质复合物。在果蝇体内，该复合物由 Drosha 蛋白和 Pasha 蛋白组成。脊椎动物体内的 Pasha 蛋白也称为 DGCR8，它含有 2 个双链 RNA 结合结构域，其作用是对 pri-miRNA 上的剪切点进行精确定位，而真正进行剪切的是 Drosha 蛋白。Drosha 有 2 个核糖核酸酶Ⅲ结构域，分

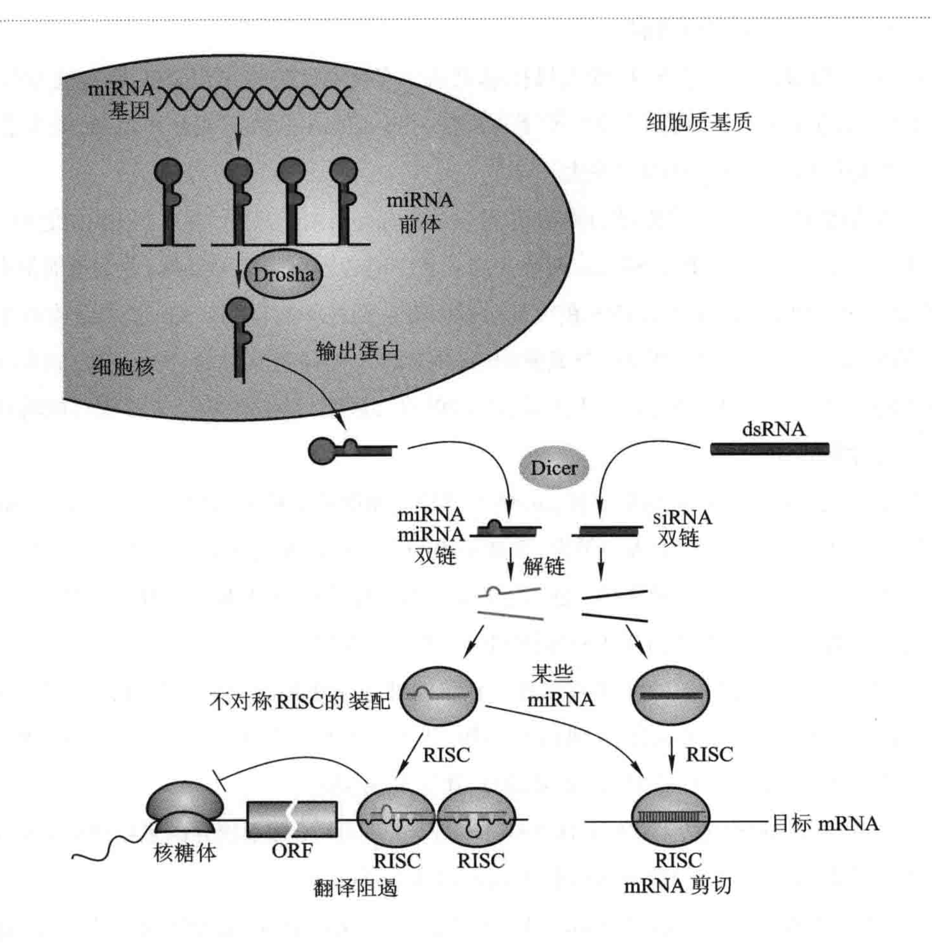


图 43-18 miRNA 和 siRNA 的产生和作用机制

别在发夹结构的茎底部离单链/双链 RNA 交界处 5′ 和 3′ 臂部 11 bp 的位置切开 RNA。

一个单一的 pri-miRNA 可能含有 1~6 个不同的 miRNAs。每一个 miRNA 都来自其中的一个发夹结构。每一个发夹结构的两翼序列是有效加工所必需的。发夹结构的双链区被 DGCR8 或 Pasha 蛋白识别，而 DGCR8 与酶 Drosha 结合，形成微加工体（microprocessor）复合物。在微加工体的作用下，每一个发夹结构被释放出来，并在 3′ 端有两个核苷酸突起。

与后生动物相比，植物 miRNA 前体上的茎环结构较大，可变性也大。此外，植物 miRNA 前体的加工只受一种在细胞核的类似切酶的酶（dicer-like1，DL1）的剪切，直接产生成熟的 miRNA，而不像后生动物，在细胞核和细胞质分别由 2 种不同的酶进行剪切。植物成熟的 miRNA 一般含有 21 nt，且 5′ 端的核苷酸多为 U，而动物则通常含有 22~23 nt。

（2）pre-miRNA 从细胞核到细胞质的运输　一旦 pri-miRNA 在细胞核完成加工，形成的 pre-miRNA 就在输出蛋白 5（EXP5）和小 G 蛋白 Ran-GTP 的作用下，离开细胞核。EXP5 能够识别 pre-miRNA3′ 端有突出（1~8 nt）的长于 14 bp 的茎结构，这种识别保证了只有被正确加工的 pre-miRNA 才能被输出。

在植物所有的 miRNA：miRNA* 双链（动物也有一亚类）输出细胞核之前，其 3′ 端突出受到一种名为 HEN1 的 RNA 甲基转移酶的催化，在核糖 2′-OH 上发生甲基化修饰，然后在 EXP5 的同源物——HST 的帮助下，从核进入细胞质。3′ 端核苷酸的甲基化修饰是植物 miRNA 行使功能所必需的，否则没有修饰的 3′ 端很容易发生多聚尿苷酸化而导致其快速降解。

（3）pre-miRNA 在细胞质中的进一步加工和成熟 miRNA 的产生　在细胞质中，pre-miRNA 在靠近末端环的位置受到切酶（dicer）的切割，释放出在 3′ 端有 2 个核苷酸突起的长约 22 nt 的 miRNA：miRNA* 双链。切酶具有多个结构域，属于依赖于 ATP 的核糖核酸酶Ⅲ家族，参与小双链 RNA 的剪切加工。有两种切酶，即切酶 1 和切酶 2，它们分别参与 miRNA 和 siRNA 的加工成熟。

5. miRNA 的功能及其作用机制

miRNA 的功能最终表现在它对特定目标基因表达的影响，它的作用首先是和通常位于目标 mRNA 3′UTR 内的互补位点结合，在此基础上再促使目标 mRNA 的降解或翻译阻遏，但少数 miRNA 也能通过上调翻译而刺激目标基因的表达。

如果是抑制翻译，miRNA 首先结合淘金者蛋白（argonaute，Ago），然后参入到 RISC 之中。Ago 是一个大的蛋白质家族，可分为 Piwi 和 Ago 两个亚族。其中前者参与转座子沉默，在生殖细胞内特别丰富，而后者通过作用于 miRNA 和 siRNA 在转录后基因表达调控中起作用。Ago 蛋白含有两个保守的 RNA 结合结构域：一个为 PAZ 结构域，与成熟的 miRNA 的 3′端单链区结合，另一个为 PIWI 结构域，作用引导链的 5′端。某些 Ago 蛋白，如人的 Ago2，能够直接切割目标 mRNA。Ago 也可能将其他蛋白招募进来，实现翻译阻遏。

在切酶 1 催化 pre-miRNA 剪切的时候，miRNA 双链开始解链。结果成熟 miRNA 只有引导链参入到 RISC 之中，与 RISC 中的 Ago 1 蛋白结合，形成 miRISC，同时过客链 miRNA* 被释放。在这里，Ago 蛋白不仅在 RISC 的形成中起关键作用，还决定了哪一条链作为引导链参入到 RISC 之中。究竟选择哪一条作为引导链主要是根据热力学不稳定性和内部茎环的位置。

与 Ago1 结合可大大提高 miRNA 的稳定性。在果蝇细胞，miRNA* 链可以和 Ago2 结合。miRNA* 是结合 Ago1 还是 Ago2，取决于 miRNA 双链的结构、热力学的稳定性和 5′端第一个核苷酸的性质。如果第一个核苷酸是 C，就与 Ago2 结合；如果是 U，就与 Ago1 结合。

其他属于 RISC 的成分包括 TRBP（HIV TAR binding protein）、干扰素诱导的蛋白质激酶激活蛋白（PACT）、SMN 复合物、脆性 X 智力低下蛋白和 Tudor-SN 蛋白等。

RISC 形成后的第一步是识别目标 mRNA。其中的 Ago1 使 miRNA 能够采取合适的方向与目标 mRNA 作用，而目标 mRNA 的确定取决于 miRNA 的引导链与目标 mRNA 上的目标序列的互补性，而互补的程度决定目标 mRNA 是水解、去稳定还是翻译阻遏。如果 miRNA 与目标 mRNA 呈现完全的互补，或者近乎完全的互补，Ago2 就直接切割 mRNA，导致 mRNA 直接降解；如果互补并不完善，就发生翻译阻遏。

识别主要通过种子序列（seed sequence）进行。种子序列是一段保守序列，多数位于 miRNA 5′端 2~7 的位置，是 miRNA 与目标 mRNA 结合所必需的。即使 miRNA 与其目标序列不完全互补，种子序列也必须与靶 mRNA 上的一段序列完全互补。

植物与后生动物的 miRNA 在识别目标 mRNA 上具有 3 个非常显著的差别：①植物 miRNA 多以转录因子的 mRNA 作为作用的对象。②植物的 miRNA 与其目标 mRNA 的序列几乎完美互补，没有或者只有很少的错配碱基，因此作用的结果将导致目标 mRNA 的降解。而后生动物的 miRNA 与目标 mRNA 的互补区横跨 5′端序列（2~7 nt）和种子序列，故对于后生动物而言，一种 miRNA 可对准同一个目标 mRNA 分子上多个位点，或者对准不同的目标 mRNA，作用的结果一般是导致翻译受阻。③后生动物的 miRNA 作用的靶点位于目标 mRNA 在 3′UTR 3 个主要的非翻译区，而植物的 miRNA 作用的靶点经常位于编码区。但不管如何，miRNA 在各种真核生物体内是高度保守的，它被认为是真核生物基因表达调控一种十分重要的机制。

RISC 对目标 mRNA 的作用可通过不同的机制进行，例如，抑制翻译的延伸、诱导核糖体的脱落和促进新生肽链的水解等，甚至去作用染色质，通过组蛋白的甲基化促进局部染色质的浓缩而导致转录受阻。RISC 还可以加快 mRNA 的去稳定，一方面促进目标 mRNA 的脱帽，另一方面促进目标 mRNA 的去尾。

一种 miRNA 的作用不一定局限在产生它的细胞，可能会扩散到其他的细胞，有时甚至在很远的地方。对于植物来说，miRNA 可以通过胞间连丝和筛管进行远距离作用。动物可能通过脂肪体进行运输，借此运输可以将一种 miRNA 特异输送到某个靶目标。有时，一种 miRNA 可在两个物种直接进行转移，

甚至发生跨界转移。例如，有人将一个针对棉铃虫基因的 miRNA 转到棉花，棉铃虫吃了棉花，miRNA 被吸收进虫体内起作用。再如，2011 年南京大学的张辰宇教授发现，植物的一种 miRNA——MIR168a 能特异性地抑制哺乳动物体内的低密度脂蛋白受体衔接蛋白 1（low density lipoprotein receptor adapter protein 1，LDLRAP1）的表达。

6. miRNA 形成的调控

编码 miRNA 的基因一般由 RNA 聚合酶Ⅱ催化转录。那些能够作用于 RNA 聚合酶Ⅱ的转录因子当然也能够影响到 miRNA 的转录。而对 miRNA 特异性的调节是编辑。通过编辑，RNA 序列上的 A 脱氨基变成 I，这种碱基转换可改变 pri-miRNA 上面的碱基配对，从而有可能阻止 Drosha 和切酶的作用。此外，编辑过的成熟 miRNA 也有可能去识别其他的目标 RNA。

7. RNAi 的应用

自 RNAi 被发现的那一刻，人们就意识到了它可能的应用前景。在后基因组时代的今天，RNAi 技术被广泛地用于研究功能基因分析、细胞信号转导通路分析、治疗各种疾病的药物设计和开发等领域（详见第四十四章“重组 DNA 技术”）。现在许多基于 RNAi 开发出来的药物已进入临床二期试验。一般而言，进行 RNAi，首先需要将合成的 siRNA 转染至细胞或动物，或者将表达 siRNA 前体——具有短发夹结构的 RNA（short-hairpin RNA，shRNA）的质粒或重组病毒引入细胞或动物。由于外源的 siRNA 与经切酶 /TRBP 剪切得来的内源 miRNA 十分相似，可被装载进入 RISC，因此也具有基因沉默的功能。同样，shRNA 和 pre-miRNA 在结构上也是类似的，在体内可以被成功地加工成成熟的 siRNA。在人工设计的时候，很容易实现让外源 siRNA 与它们的靶 mRNA 完全互补，从而直接导致靶 mRNA 的降解。

框 43-2　生化研究动态——我们吃的不仅是食物，也许还有“信息”

根据 2011 年 9 月 22 日 *Cell Research* 的一篇题为“Exogenous plant MIR168a specifically targets mammalian LDLRAP1: evidence of cross-kingdom regulation by microRNA”的论文，南京大学生命科学学院的张辰宇教授等人发现，植物的微 RNA（miRNA）可以通过日常食物摄取的方式，进入人体血液和组织器官，并通过调控人体内靶基因表达的方式，影响人体的生理功能，进而发挥生物学作用。

miRNA 是一类非编码小 RNA（参看第四十三章“真核生物的基因表达调控”），它通过与靶基因的 mRNA 结合的方式，抑制相应的蛋白质翻译。张辰宇之前的研究成果已表明，某些 miRNA 是由组织和细胞主动分泌的（分泌型 miRNA），可稳定存在于哺乳动物的血清和血浆中。因此，miRNA 是一类新型的疾病标志物，可应用于疾病（如肿瘤）的早期诊断。分泌型 miRNA 也是一类新型信号分子，调控细胞间和组织间的信号转导。

在本项研究中，张辰宇等发现，外源性的植物 miRNA 可以在多种动物的血清和组织内检测到，并且它们主要是通过饮食的方式摄入到体内。其中编号为 168a 的植物 miRNA（miR-168a）富含于稻米中，是中国人血清中含量最丰富的一种 miRNA。体内和体外的研究表明，植物 miR-168a 可以结合人和小鼠的低密度脂蛋白受体衔接蛋白 1（low density lipoprotein receptor adapter protein 1）的 mRNA，从而抑制其在肝脏的表达，进而减缓低密度脂蛋白从血浆中的清除。这些发现表明，食物中的外源性植物 miRNA 可以通过调控哺乳动物体内靶基因表达的方式，影响摄食者的生理功能。

该发现引起了世界顶级学术刊物和主流媒体的普遍关注，其意义重大，至少提示我们：人类在从外界摄取食物的时候，不仅在吃“食物”其中的六大营养物质，也许还在摄入“信息”。该发现从一种新的角度诠释了“吃什么补什么”（You are what you eat）。此外，对理解生物之间尤其跨“界”（比如在动植物间）的相互作用，以及共进化（co-evolution）等提供了新的线索。当然对于中国人来说，还为研究传统中草药的作用方式提供了新的思路。

8. RNAi的生物学意义

目前普遍认为，RNAi在植物和昆虫体内相当于一种免疫系统，起着保卫基因组的作用，它能够防止外来的有害基因或病毒基因整合到植物基因组中。因为许多病毒的基因组为双链RNA，或者在复制过程中经历双链RNA中间体，这些双链RNA可被宿主细胞内的切酶和RISC识别并切割而失去活性。此外，它还参与基因表达的调控。目前，人们已经鉴别出1 000多个人类miRNA，它们参与细胞分化、增殖、凋亡、胰岛素分泌以及心脏、大脑和骨骼肌的发育过程。已发现某些miRNA在某些生物胚胎发育的特定时段起作用，参与调控发育时钟，因此这些miRNA有时被称为小时序RNA(stRNA，small temporal RNA)。

五、mRNA的稳定性与基因表达调控

细胞质中所有的RNA都会水解，但是不同类型的RNA在细胞内的半衰期并不相同。显然一种mRNA半衰期越高，其被翻译的机会就越高。当一种mRNA的稳定性能受到特定的调节信号作用而发生变化的时刻，mRNA的稳定性和降解便成为调节基因表达的一种重要手段。

已发现一系列的序列和二级结构元件能够影响mRNA的稳定性，例如许多“短命”mRNA在3′UTR有一种富含AU序列的元件(AU-rich element，ARE)。ARE的作用一方面可提高mRNA的脱尾作用而降低mRNA的稳定性，另一方面可直接参与翻译阻遏。

铁是一种必需元素，作为细胞内多种蛋白质或酶的组分，但细胞内过量的铁会诱发有害的自由基反应，导致脂质、核酸和蛋白质的损伤。因此，细胞内的铁浓度是受到严格调控的。

细胞内对铁浓度的控制是通过对2种参与铁代谢的蛋白质——转铁蛋白受体(transferrin receptor，TfR)和铁蛋白(ferritin)的翻译而实现的。铁蛋白的功能是贮存细胞内的铁，而转铁蛋白受体的功能是通过与血液内运输铁的蛋白质——转铁蛋白的相互作用，调节进入细胞的铁的量。如果细胞内铁浓度较高，铁蛋白的翻译就上调，TfR的翻译就下调，以防止细胞因摄入过多的铁而中毒；反之，如果细胞内铁浓度较低，铁蛋白的翻译就下调，TfR的翻译就上调，以满足细胞对铁的需要。

如图43-19所示，由茎环结构构成的铁应答元件(iron-response element，IRE)分别在TfR 3′UTR

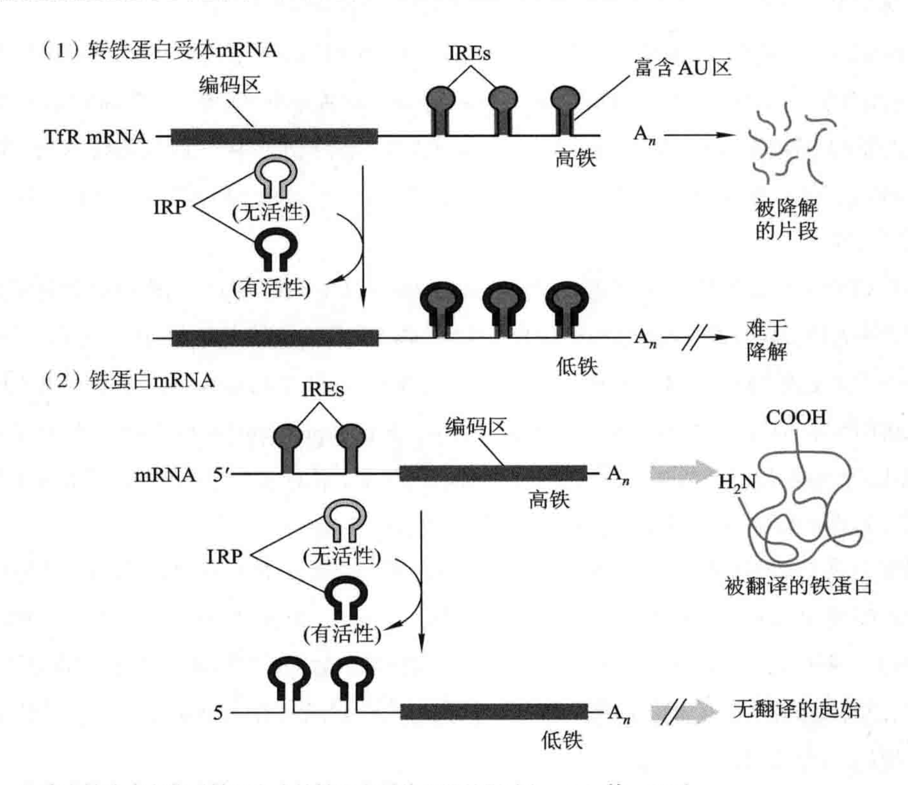


图43-19 细胞内铁浓度变化对铁蛋白或转铁蛋白受体翻译的影响(Lodish等，2007)

和铁蛋白的 5′UTR 内，而一种 IRE 结合蛋白（IRE-binding protein，IREBP）通过与 IRE 的结合，可以调节上述两种 mRNA 的稳定性，其中 TfR 上的 IRE 含有去稳定的 ARE。当细胞处于低铁状态下，IREBP 有活性，这时 IREBP 与 IRE 结合，对 TfR mRNA 和铁蛋白 mRNA 的稳定性和可翻译性产生不同的影响，由于铁蛋白 mRNA 上的 IRE 位于 5′UTR，因此将会阻止翻译，相反，TfR mRNA 上的 IRE 位于 3′UTR，因此 IREBP 与其结合将提高 TfR mRNA 的稳定性；当细胞处于高铁状态下，IREBP 无活性，这时 IREBP 无法与 IRE 结合，TfR mRNA 因失去 IREBP 的保护而被水解，而铁蛋白 mRNA 的 IRE 因没有结合 IREBP 而被翻译。

Quiz13 你如何设计一个实验，证明 TfR 和铁蛋白的基因表达调控是在翻译水平，而不是在转录水平上进行的？

六、对翻译过程本身的调节

1. 对翻译起始阶段的调节　此过程主要是通过对两种起始因子 eIF4E 和 eIF2 的磷酸化修饰而实现的。细胞内某些信号（生长因子、热激、病毒感染、有丝分裂和血红素浓度变化等）能够诱发这两种起始因子的磷酸化或去磷酸化，从而改变翻译的效率。此外，细胞内的某些抑制因子能够与特定的起始因子结合，而抑制翻译的起始。eIF4E 和 eIF2 的磷酸化对翻译的影响正好相反，前者的磷酸化可刺激翻译，后者则抑制翻译。

（2）对翻译终止阶段的调节　此过程是通过再次程序化的遗传解码而进行的（参看第四十一章“再次程序化的遗传解码和翻译暂停”）。

有关翻译后水平的调节的内容请参看第四十章“蛋白质的翻译后加工及其定向和分拣”。

科学故事——RNAi 的发现

早在 1990 年，Rich Jorgensen 等人在对矮牵牛（petunias）进行转基因研究中发现一个奇怪的现象：将一个能产生色素的基因——蝴蝶兰查尔酮合酶（chalone synthase）基因置于一个强启动子后，引入矮脚牵牛，以加深花朵的紫色，结果非但没有看到期待中的深紫色花朵，倒是多数花成了花斑，甚至变为白色。Jorgensen 将这种现象命名为共阻遏（cosuppression），因为他认为导入的基因和其机体内相似的内源基因同时都被阻遏而不表达。与此类似的实验是，将经基因改造过的含有一个植物基因片段的 RNA 病毒感染植物细胞，也可以导致内源植物基因的沉默。开始人们认为，共阻遏是矮牵牛特有的怪现象，但后来发现在其他许多植物中，甚至在真菌中也有类似的现象。

那么是什么原因导致共阻遏现象发生的呢？对此，有人认为转基因引发的共阻遏现象可能是其诱导了基因发生特异的甲基化，导致基因不能转录，这种阻遏称为转录水平基因沉默（transcriptional gene silencing，TGS）。但也有人认为，基因沉默是在转录后发生的，称为转录后基因沉默（post-transcriptional gene silencing，PTGS）。后来的核转移实验表明，内源基因的转录并没有停止，只是转录物进入胞浆后很快被降解，没有积累，这说明 Jorgensen 发现的共阻遏属于 PTGS。后来进一步证实，在植物中，将出现转基因导致基因沉默的植物嫁接到另一没有基因沉默的植物中，同样可以诱发 PTGS。

到了 1995 年，美国康乃尔大学的 Su Guo 和 Kemphues，在将 *par-1* 基因的反义 RNA 链注射到线虫的生殖腺后，也发现了一个非常奇怪的现象：他们本来想利用反义 RNA 技术，特异性阻断 *par-1* 基因的表达，同时亦在其对照组实验中，给线虫注射正义 RNA（sense RNA）以期观察到 *par-1* 基因表达的增强。但得到的结果竟是，二者以同样的高频率诱导 *par-1* 的缺失表型（null phenotype）的出现。这与传统上对反义 RNA 技术的解释竟是正好相反。研究小组一直没能对这个意外结果予以合理的解释。

Guo 和 Kemphues 实验留下来的谜团直到 1998 年才由 Andrew Fire 等人解开。通过大量艰苦的工作，他们证实，Su Guo 和 Kemphues 遇到的正义 RNA 抑制基因表达的现象，以及过去的反义 RNA 技术对基因表达的阻断，都是由于体外转录所得 RNA 中污染了微量双链 RNA 而引起。如果仅仅注射经高度纯化后的目标基因的单链 RNA，抑制效应就变得十分微弱；此外，使用与目标基因无关的双链 RNA 无任何

影响。Fire 随后的研究表明，双链 RNA 对目标基因表达所产生的效应是因为它们影响到了 mRNA 的稳定性，而且，双链 RNA 似乎能够越过细胞间的障碍，在注射点以外的地方起作用，更让人吃惊的是，每个细胞只要有几个分子的双链 RNA 就可以起作用，这一点意味着在其中有催化或放大机制。进一步研究还表明，只有对应于目标基因外显子序列的双链 RNA 才有上述效应，与内含子序列对应的双链 RNA 无效。后来的实验表明，在线虫中注入双链 RNA 不仅可以阻断整个线虫的同源基因表达，还会导致其子一代的同源基因沉默，该小组将这一现象称为 RNA 干扰（RNA interference，RNAi）。Fire 也因此荣获 2006 年诺贝尔生理学或医学奖。

到目前为止，RNAi 现象被广泛地发现在真菌、拟南芥、水螅、涡虫、锥虫、果蝇、斑马鱼等许多真核生物中。这说明 RNAi 很可能出现在生命进化的早期阶段。随着研究的不断深入，RNAi 的机制正逐步地被阐明，同时亦成为功能基因组研究领域中的有力工具，用于鉴定功能缺失表型。RNAi 也越来越为人们所重视。

种种迹象表明，RNA 干扰是真核生物体内的一种相当普遍的现象，难怪 *Science* 将 2002 年的年度分子授予小 RNA 和 RNAi。

e43-3 本章小结填填看

推荐网址：

1. http://www.emunix.emich.edu/~rwinning/genetics/eureg.htm（美国东密歇根大学一个遗传学网站，有关真核基因表达调控的内容）
2. http://www.emc.maricopa.edu/faculty/farabee/biobk/biobookgenctrl.html（美国马里科帕郡社区学院一个在线的分子遗传学教程，有关基因表达调控的内容）
3. http://www.uic.edu/classes/bios/bios100/lecturesf04am/lect15.htm（美国加利福尼亚大学欧文分校一个生物学网站，有关含基因表达调控的内容）
4. http://en.wikipedia.org/wiki/RNA_interference（维基百科有关 RNA 干扰的内容）

参考文献：

1. Terns M P Terns R M. CRISPR-based adaptive immune systems. *Current Opinion in Microbiology*, 4(3): 321–327, 2011.
2. Singer R H, et al. Single-molecule mRNA decay measurements reveal promoter- regulated mRNA stability in yeast. *Cell*, 147(7): 1 484–1 497, 2011.
3. Hale C R, et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas Protein Complex. *Cell*, 139(5): 945–956, 2009.
4. Madhani H D, et al. Histone variant H2A.Z marks the 5′ends of both active and inactive genes in Euchromatin. *Cell*, 123: 233–248, 2005.
5. Wilhelm B T. Complex controls: the role of alternative promoters in mammalian genomes. *Trends in Genetics*, 19(11): 640–648, 2003.
6. Martens J A, Winston F. Recent advances in understanding chromatin remodeling by Swi/Snf complexes. *Curr Opin Genet Dev*, 13: 136–142, 2003.
7. Hannon G J. RNA interference. *Nature*, 418: 244–251, 2002.
8. Fire A, et al. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, 391: 806–811, 1998.

第四十四章　重组 DNA 技术

在 1970 年以前，以核酸为对象的研究者面对一堆不均一的 DNA 样品往往束手无策，这是因为当时他们无法从中大量制备一种单一序列的 DNA 片段。由于大多数的基因在二倍体细胞是双拷贝，而在单倍体细胞为单拷贝，因此，想从一种生物体内直接分离得到一个基因几乎是不可能的事情。然而，诞生于 20 世纪 70 年代初的重组 DNA 技术不但使分离一个基因不再困难，而且借助这项技术，还可以在不同的系统中对其进行表达和改造。如今，重组 DNA 技术和由它衍生出来的其他技术已成为生命科学研究不可缺少的利器，每一个从事生命科学研究的人都应该对它们有所了解，故本章就集中介绍这些技术的原理和应用。

第一节　重组 DNA 技术简介

重组 DNA 顾名思义就是通过某种手段，将不同来源的 DNA 片段连接起来，并进行扩增和纯化以供进一步研究的技术（图 44-1）。

当克隆（无性繁殖）这一名词被借用到同一种重组 DNA 分子的大量扩增和纯化的时候，基因克隆或分子克隆等术语便应运而生，因此重组 DNA 又被称为基因克隆（genetic cloning）或分子克隆（molecular cloning）。

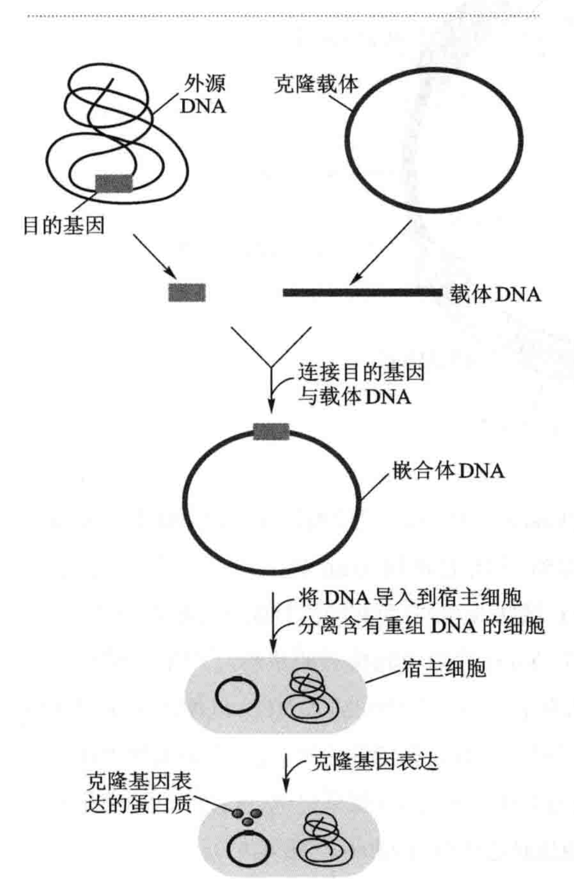


图 44-1　重组 DNA 技术的基本步骤

有效的基因克隆至少需要满足 5 个条件：①具有容纳外源基因或序列的载体（vector）；②具有将外源基因或序列导入到载体的工具；③具有合适的宿主细胞；④具有将重组 DNA 引入到宿主细胞的有效途径；⑤具有选择和筛选重组体的方法。

一、基因克隆的载体

载体的作用是容纳被克隆的目的基因，以便将目的基因带入到特定的宿主细胞进行扩增或表达。一种理想的载体一般需要满足以下几个条件：①含有原核生物 DNA 复制起始区，以便于在原核系统中的扩增。②含有真核生物 DNA 复制起始区，以方便在真核细胞内的自主复制。③含有集中了多种常用的限制性内切酶切点的多克隆位点（multiple cloning site，MCS），以方便各种克隆片段的插入。④带有抗生素抗性基因或其他选择性标记，以方便克隆的筛选和鉴别。⑤含有可诱导的或组织特异性的启动子或增强子序列，有利于控制被插入的基因在宿主细胞内的表达。⑥现代的载体一般含有多功能的结构元件，同时兼顾到克隆、DNA 测序、体外突变、转录和自主复制。

Quiz1　你如何设计一个实验，能克隆到大肠杆菌的 DNA 复制起始区？

目前使用的载体多衍生于质粒、噬菌体和病毒。市场上提供了种类繁多的商业化载体，而选择哪一种载体，取决于实验系统的设计、如何筛选以及如何利用克隆的基因。

（一）质粒载体

质粒（plasmid）主要是指细菌染色体以外的、能自主复制并与细菌共生的遗传成分。其特点包括：①是一种共价闭环双链 DNA，可形成天然的超螺旋结构。不同质粒的大小在 2~300 kb，小于 15 kb 的小质粒比较容易分离纯化，且能容纳较大的外源 DNA，因此最适合用做载体。②含有 DNA 复制起始区，因而能自主复制。按复制的调控机制及其拷贝数，可将它们分为 2 类：一类为严紧控制（stringent control）型，其复制受到严格的控制，拷贝数较少，只有一到几十个；另一类是松弛控制（relaxed control）型，其复制不受宿主细胞控制，每个细胞有几十到几百个拷贝。显然，松弛型质粒更适合作为克隆载体。③对宿主细胞的生存不是必需的，但通常带有有利于宿主菌在特定条件下生存的基因。例如，许多天然的质粒（R 质粒）带有抗药性基因，能编码某种酶分解或破坏常见的抗生素。

Quiz2 如果宿主菌是古菌，这些抗生素抗性标记还有效吗？为什么？

基因克隆中使用的质粒载体基本上都是经改造过的松弛型质粒，其内部一些无用的序列已被去除，同时引进了一些有用的序列。最常用的大肠杆菌克隆质粒为 pUC18/19（图 44–2），它是由天然的 pBR322 质粒改造而来。此质粒的复制起始区能高频启动自身的复制，使一个细菌细胞内的质粒拷贝数高达 500 个以上。pUC18 和 pUC19 的差别仅仅在于 MCS 的方向正好相反。此外，该质粒还携带一个抗氨苄青霉素基因（*amp*^{R} 或 *Ap*^{R}），由它编码的一种内酰胺酶（β-lactamase，bla），能打开青霉素分子内的β- 内酰胺环，使氨苄青霉素失效。因此，当细菌用 pUC18/19 转化后，放入含氨苄青霉素的培养基中，凡不含 pUC18/19 者都不能生长，而长出的细菌一定带有 pUC18/19。

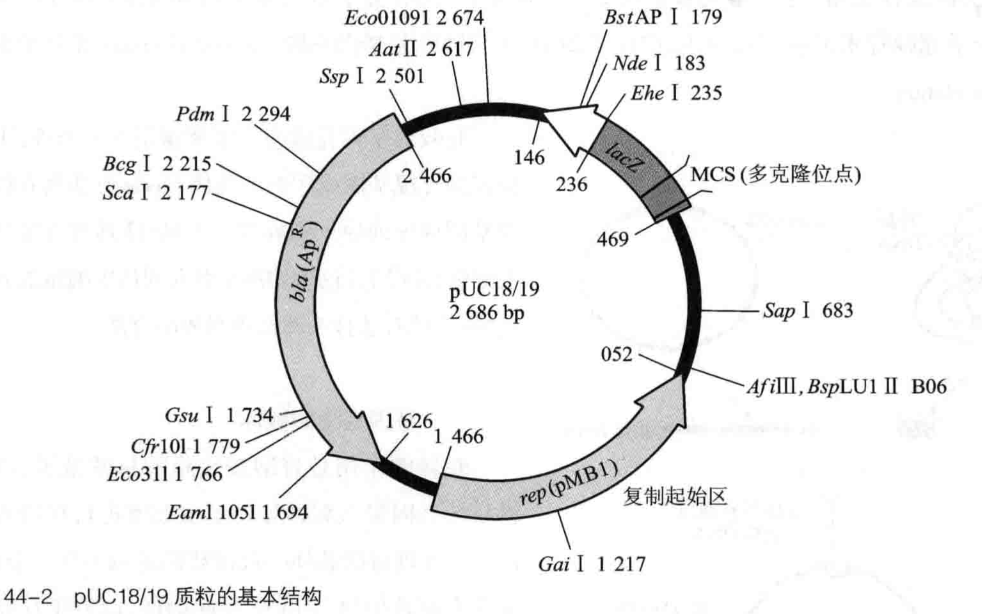


图 44–2　pUC18/19 质粒的基本结构

pUC18/19 还携带大肠杆菌乳糖操纵子的 *lacI* 和 *lacZ'*，但 *lacZ'* 仅仅编码β- 半乳糖苷酶 N 端的 146 个氨基酸残基（α 肽）。当培养基中含有诱导物 IPTG 和显色底物 X-gal 时，*lacZ'* 被诱导表达，表达出来的β- 半乳糖苷酶 N 端肽段能与宿主菌表达的 C 端肽段互补，而共同组装成具有活性的β- 半乳糖苷酶，此现象被称为 α- 互补（α-complementation）。X-gal 受到半乳糖苷酶水解后产生有颜色的产物，从而使菌落呈现蓝色。通常在不改变 ORF 的前提下，在 *lacZ'* 中间插入 MCS，以便外来序列的插入。当外来序列插入后，可打破 *lacZ'* 原来的 ORF，致使半乳糖苷酶活性丧失，这种现象被称为插入失活（insertional inactivation）。含有重组体质粒的菌落因无法水解 X-gal 就呈白色，这种颜色的变化经常用来筛选含有重组质粒的转化菌落，此法被称为蓝白斑筛选法（blue/white）（图 44–3）。

（二）噬菌体载体

噬菌体载体多由 λ 噬菌体改造而成，它们常用来克隆较大的 DNA 片段，特别是构建真核生物的

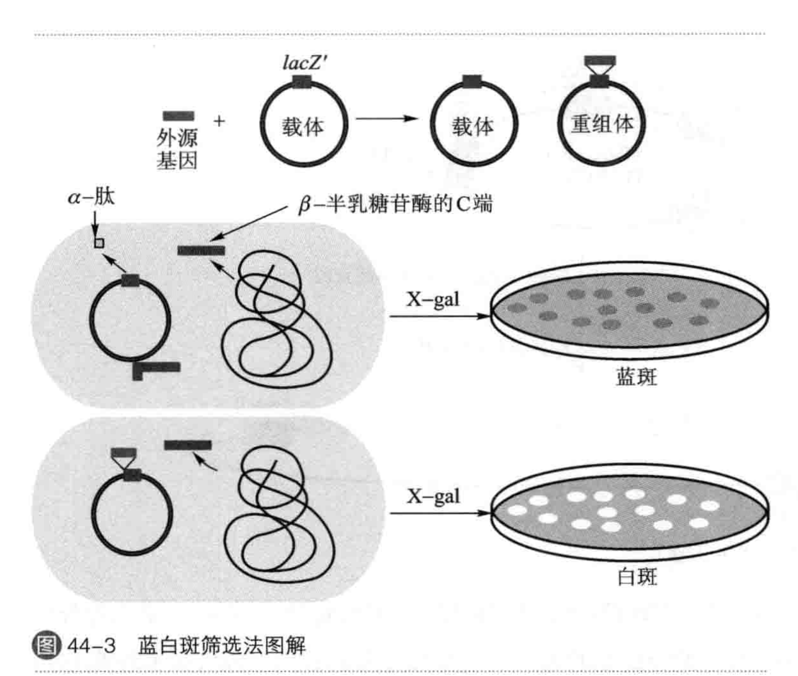


图 44-3 蓝白斑筛选法图解

cDNA 文库(cDNA library)或基因组文库(genomic library)。

利用λ噬菌体载体,首先需要用外源 DNA 替代噬菌体 DNA 的中段非必需序列,或者将其插入到噬菌体 DNA 的中段,然后让重组的 DNA 随噬菌体的左右臂一起,在体外包装成噬菌体,再让噬菌体去感染大肠杆菌(图 44-4)。现在使用的λ噬菌体载体已做了以下几个方面的改造:①去除了λ-DNA 上的一些限制性酶切点。②在中部非必需区域,替换或插入 MCS 和某些标记基因,如可供蓝白斑筛选的 *lac*I-*lac*Z′ 等基因。由此可构建出插入型载体(insertion vector)和替换型载体(substitution vector)。插入或置换中段的外源 DNA 长度是有一定限制的,太长或太短都会影响到包装后的重组噬菌体的生存活力。

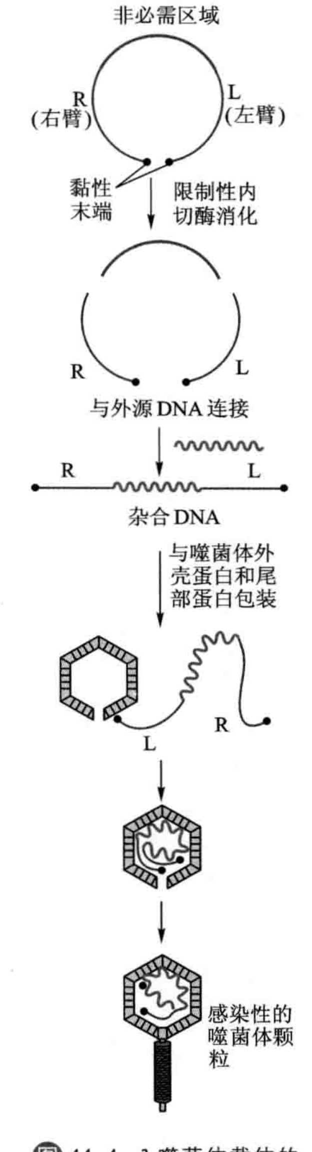


图 44-4 λ噬菌体载体的构建、重组和包装

使用噬菌体载体至少有两个好处:一是能够容纳较长的外源 DNA,二是其感染宿主菌要比质粒转化细菌的效率高得多。其缺点则在于克隆操作要比质粒载体繁琐。

（三）黏粒(cosmid)

这是一种兼有部分λ-DNA 和部分质粒序列的杂合载体。其中,来自λ-DNA 的成分是噬菌体体外包装所必需的 *cos* 序列,来自质粒的成分包括复制起始区、特定的抗生素抗性基因和 MCS。

在使用黏粒的时候,重组体 DNA 在体外与野生型λ噬菌体的外壳蛋白和尾部蛋白包装成感染性的颗粒,以便能高效进入宿主细胞。在宿主细胞内,它们又像质粒一样进行复制,但由于缺乏外壳蛋白的基因,因此并不能在宿主细胞内形成新的噬菌体颗粒。

黏粒可插入长 30~45 kb 的外源 DNA,对 DNA 的大小有选择作用,主要用于 DNA 文库的构建。

（四）PAC 和 BAC

PAC 即是 P1 噬菌体人工染色体(P1 phage artificial chromosome),能容纳 100~300 kb 的外源 DNA。与λ噬菌体相似,P1 噬菌体在大肠杆菌内以原噬菌体的形式存在,但它并不整合到大肠杆菌染色体上。

BAC 是指细菌人工染色体(bacterial artificial chromosome)。它是一种非常稳定的质粒载体,能容纳的外源 DNA 在 100~250 kb,最长的可达 1 Mb。BAC 含有严紧型质粒 F 因子的复制起始区(*oriF*),因此其拷贝数受到严格的控制(1~2 个拷贝 / 细胞)。此外,它还含有 MCS 和选择性标记基因 cam^R 或 *sacB*Ⅱ,以及能驱动基因转录的启动子。cam^R 为氯霉素抗性基因。*sacB*Ⅱ编码的是果聚糖蔗糖酶(levansucrase),该酶能将蔗糖转化成果聚糖。果聚糖对细菌是有毒的,因此,如果 *Sac*BⅡ有活性,细菌就无法生存。当外源 DNA 插入到 *Sac*BⅡ的上游以后,*Sac*BⅡ就无法正常转录。利用此性质,可进行正选择重组体(图 44-5)。

（五）YAC 和 HAC

YAC 是指酵母人工染色体(yeast artificial chromosome),是由酵母、原生动物和细菌质粒 DNA 组建成的杂合载体,能够容纳大于 1 Mb 的外源 DNA。其中的自主复制序列、选择性标记基因和着丝点来自酵母,端粒序列一般来自四膜虫,MCS 和原核复制起始区来自质粒。在酵母细胞内,YAC 像酵母染色体一样行使功能,而在大肠杆菌内,又能像质粒一样复制。

YAC 上的标记基因有:青霉素抗性基因(*bla*)——用于转化大肠杆菌时的选择;*TRP1*——参与色

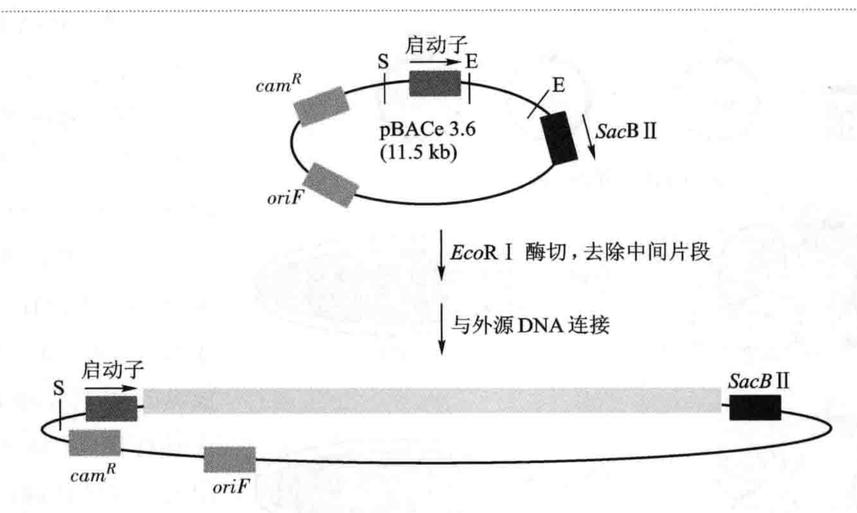


图 44-5 细菌人工染色体的结构和外源 DNA 的插入

Quiz3 赭石型校正 tRNA 能校正哪一种无义突变?

氨酸合成的一个关键酶的基因，为色氨酸营养缺陷型酵母细胞提供选择性标记；*SUP4*——赭石型校正 tRNA（ochre suppressor tRNA）基因，赋予酵母细胞的无义突变株具有野生表型，但是，如果外源 DNA 插入到 *SUP4* 内部导致其失活，那宿主酵母就能够维持突变表型，此性质可用来筛选重组体。

YAC 克隆的基本策略是（图 44-6）：首先使用限制性内切酶 *Bam*HⅠ（图中的 B 为它的切点）消化载体，以便游离出端粒，产生线形染色体结构；然后，使用 *Sma*Ⅰ（图中的 S 为它的切点）分别消化外源 DNA 和载体；最后，使用连接酶将目标 DNA 插入到 *SUP4* 内部，再转化宿主细胞，并进行筛选。

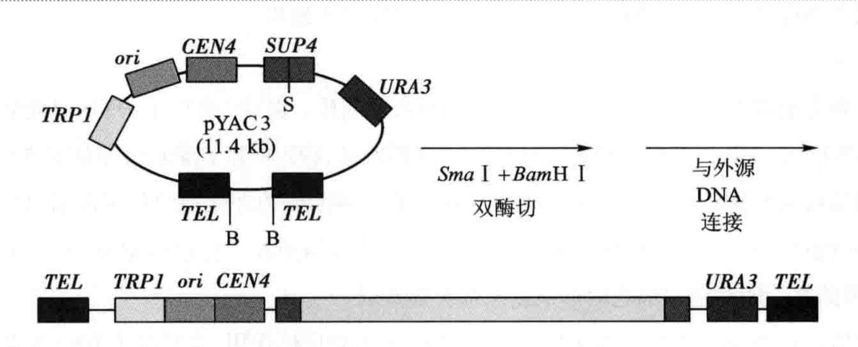


图 44-6 酵母人工染色体的结构和外源 DNA 的插入

HAC 是指人人工染色体（Human artificial chromosome）。它衍生于 BAC、YAC 或 PAC，含有自主复制序列、端粒和 α-卫星 DNA(着丝点 DNA)，能够容纳约 100 kb 的外源 DNA，在宿主细胞内不发生整合，可在有丝分裂期间稳定地传给子细胞。

（六）真核细胞病毒载体

这是一类由动、植物病毒改造而成的载体。由于动物细胞的培养和操作较复杂、花费也较大，在动物病毒载体构建时，一般都把质粒的复制起始区放置其中，形成穿梭载体（shuttle vector），以便能在细菌中大量扩增，再引入到真核细胞。使用动物病毒载体的目的多为将目的基因或序列引入动物细胞中表达，或者测试其功能，或作为基因治疗的载体。

二、将外源基因或序列导入载体的工具

将外源基因或序列导入到载体需要特殊的工具酶，其中以限制性内切酶（restriction endonuclease，RE）和连接酶最为重要，此外，有时还需要 DNA 聚合酶、多聚核苷酸激酶和 S1 核酸酶等，这里主要介绍限制性内切酶。

（一）限制性内切酶

重组 DNA 首先需要对特定的 DNA 进行精确的定向切割，RE 的发现使之成为可能。RE 实际上

是一类特殊的具有高度位点特异性的 DNA 内切酶，可识别双链 DNA 分子内部特殊的碱基序列（通常为 4~6 bp 的回文序列），切开 DNA 的两条链，产生特定的末端。

到目前为止，已发现 4 类 RE：①Ⅰ类是由 3 种不同的亚基组成，兼有甲基化酶和依赖于 ATP 的内切酶活性，能识别 DNA 分子特定的碱基序列，同时随机切断在识别位点以外的 DNA 序列，这类酶的作用需要 Mg^{2+}、SAM 及 ATP；②Ⅱ类不具有修饰酶活性，只由一条肽链组成，需要 Mg^{2+}，但不需要 SAM 和 ATP，其切割 DNA 特异性最强，且在识别位点内部切断 DNA，93% 的 RE 属于此类；③Ⅲ类与Ⅰ类相似，需要 Mg^{2+} 和 ATP，但切点在识别序列周围 25~30 bp 范围内，显然，Ⅱ类最适合于基因克隆，在重组 DNA 技术中提到的 RE 通常都属于此类；④Ⅳ类能切割甲基化位点，例如 *Msp* Ⅰ识别的序列是 CCGG，不管 C 是否甲基化都能够切割。

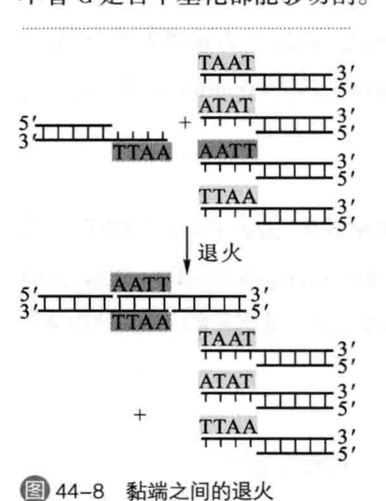


图 44-8　黏端之间的退火

所有 RE 切割 DNA 后在切点产生的总是 5′－磷酸和 3′－OH，这对于后面的连接反应十分重要，原因是 DNA 连接酶要求连接点必须是 5′－磷酸和 3′－OH。然而，不同的 RE 在切割 DNA 的时候，切割的方式不一定相同（图 44-7）。根据在识别位点上的切割方式，Ⅱ类限制性内切酶又分为 2 个亚类：第一亚类交错切开 DNA 的两条链，产生突出的互补末端（有的产生 5′突出，有的产生 3′突出），这样的末端在特定的条件下，能够重新缔合在一起，因此被称为黏端（cohesive end）；第二亚类在 DNA 两条链相同的位置切开 DNA，产生无突出的平端（blunt end）。在基因克隆中，使用最多的是产生黏端的 RE，因为不同的 DNA 分子经过同一种 RE 处理后，产生相同的黏端，经退火后很容易“粘”在一起，从而大大方便了后面的连接反应（图 44-8）。

Hpa Ⅰ
5′-G-T-T-A-A-C-3′
3′-C-A-A-T-T-G-5′
切点

*Eco*R Ⅰ
5′-G-A-A-T-T-C-3′
3′-C-T-T-A-A-G-5′
切点

Hind Ⅲ
5′-C-T-G-C-A-G-3′
3′-G-A-C-G-T-C-5′
切点　*Pst* Ⅰ

图 44-7　Ⅱ类限制性内切酶的三种切割方式

不同的 RE 识别的 DNA 序列长度可能并不相同，识别序列有 4 bp、6 bp 或 8 bp，其中以识别 6 bp 的最为常见（表 44-1）。显然，一个由随机碱基序列组成的 DNA，如果受到识别 4 bp、6 bp 或 8 bp 的 RE 的消化，那大概分别平均每隔 256 bp（4^4）、4 096 bp（4^6）和 65 536 bp（4^8）就产生一个切点。当需要对一个已知长度的长 DNA 分子进行切割的时候，可按此计算，粗略估计出它对不同 RE 可能的切点频率，以便选择合适的 RE。

表 44-1　常见的几种 RE 识别的碱基序列和切点性质

限制性内切酶	识别序列和切点	限制性内切酶	识别序列和切点
Alu Ⅰ	AG↓CT	*Hpa* Ⅱ	CC↓GG
*Bam*H Ⅰ	G↓GATCC	*Kpn* Ⅰ	GGTAC↓C
Bgl Ⅱ	A↓GATCT	*Mbo* Ⅰ	↓GATC
*Eco*R Ⅰ	G↓AATTC	*Pst* Ⅰ	CTGCA↓G
Hae Ⅲ	GG↓CC	*Sma* Ⅰ	CCC↓GGG
Hind Ⅲ	A↓AGCTT	*Not* Ⅰ	GC↓GGCCGC

（二）DNA 连接酶

连接酶在基因克隆中的作用是将外源 DNA 连接到载体上，使用最多的是 T4 噬菌体编码的 DNA 连接酶，它以 ATP 为能源，不仅能连接黏端 DNA，还能连接平端 DNA。

（三）DNA 聚合酶

在基因克隆中使用的 DNA 聚合酶有 Klenow 酶、T4 DNA 聚合酶和以 *Taq* DNA 聚合酶为代表的耐热性 DNA 聚合酶。

DNA 聚合酶主要用于：①对 3′端隐缩的 DNA 进行填补或末端标记；②合成 cDNA 的第二条链（参见 cDNA 文库的建立）；③利用缺口平移，制备 DNA 探针；④末端终止法测定 DNA 序列；⑤PCR（参看本章第四节“聚合酶链式反应”）。

e44-1 缺口平移

(四) 逆转录酶

逆转录酶主要用于以 mRNA 为模板，合成 cDNA。基因克隆中经常使用的逆转录酶来自两种逆转录病毒——禽类成髓细胞瘤病毒(AMV)和莫洛尼鼠白血病病毒(MMLV)。

三、宿主细胞

宿主细胞是接受、扩增和表达重组 DNA 的场所。理论上，任何活细胞都可以作为宿主细胞，但最为常用的有大肠杆菌、酵母和哺乳动物的培养细胞等。

四、将重组 DNA 引入到宿主细胞的途径

目的基因序列与载体连接后，要导入细胞中才能复制、扩增，再经过筛选，才能获得重组 DNA 分子克隆。将重组体引入到宿主细胞的主要方法包括:转化(transformation)、转染(transfection)及电穿孔(electroporation)、脂质体介导和弹道基因转移。

(一) 转化

基因克隆中的转化是指质粒进入宿主细胞的过程。为了提高转化效率，需要采取一些特殊方法处理细胞，经处理后的细胞更容易接受外源 DNA，因此被称为感受态细胞(competent cell)。例如，大肠杆菌经冰冷的 $CaCl_2$ 处理，其表面通透性增加，就成为感受态细菌。此时加入重组质粒，并突然由 4℃ 转入 42℃做短时间热休克处理，质粒就能顺利进入细菌。

(二) 转染

重组的噬菌体 DNA 可像质粒 DNA 一样进入感受态细菌，这种方式被称为转染。

重组 DNA 进入哺乳动物细胞也称为转染。常用的是 DNA- 磷酸钙共沉淀法，其原理是:DNA 在以磷酸钙 -DNA 共沉淀物形式出现时，培养细胞摄取 DNA 的效率会显著提高。

(三) 电穿孔

用高压脉冲短暂作用于细菌也能显著提高转化效率，这种方法称为电穿孔。用电穿孔法处理培养的哺乳动物细胞，可使细胞瞬间出现可逆性穿孔，从而使 DNA 导入到胞内，但所用外加电场强度和电脉冲的长度等条件与处理细菌有较大的差别。

(四) 脂质体介导

近年来用人工脂质体包装 DNA，形成的脂质体通过与宿主细胞的细胞膜融合，而将 DNA 导入细胞。此方法简单而有效，现有各种商业化的脂质体试剂可供使用。

(五) 弹道基因转移

Quiz4 以上哪些方法不能直接用于向完整的植物细胞导入外源 DNA？

弹道基因转移(ballistic gene transfer)使用细小的由 DNA 包被的特制子弹(DNA-coated projectile)作为载体，在基因枪(gene gun)的帮助下，将重组 DNA “射入”到受体细胞。

五、重组体的选择和筛选

外源 DNA 与载体正确连接的效率，以及重组体导入宿主细胞的效率都不可能是百分之百的，因此，最后培养出来的细胞群中只有小部分是含有外源 DNA 的重组体。只有把含有目的重组体的宿主细胞从各种无关的细胞中筛选出来，才等于获得了目的DNA的克隆。因此，筛选是基因克隆不可缺少的一步。

筛选方法一般可分为直接筛选和间接筛选，前者根据宿主细胞接受外源基因以后直接引起表型的变化而进行筛选。然而，多数外源 DNA 没有可利用的表型，因而需要使用后一种方法，来对重组体 DNA 序列和表达产物进行分析。

(一) 直接筛选

可根据抗生素敏感性和抗性变化进行的筛选(参看本章第一节)，也可根据营养缺陷型的恢复以及蓝白斑筛选(参看本章第一节)。直接筛选可以排除大量的非目的重组体，但这还只是粗筛。例如，细菌

e44-2 营养缺陷型恢复的筛选

可能发生变异而引起抗药性的改变，并不代表外源序列的插入，因此仍需要做进一步细致的筛选。

（二）间接筛选

（1）核酸杂交法　利用标记的核酸做探针，与转化细胞的DNA或RNA进行杂交，可以直接筛选和鉴定目的序列克隆。其中以DNA为杂交对象的杂交法称为Southern杂交或Southern印迹，以RNA为杂交对象的方法称为Northern杂交或Northern印迹。核酸探针可以用放射性同位素标记，也可以用非放射性物质标记，前者用放射自显影显示阳性克隆，后者借助于特殊的显色反应显示阳性克隆。

（2）PCR法　PCR技术的出现给克隆的筛选增加了一个新手段。如果已知目的序列的长度和两端的序列，就可以设计合成一对引物，以转化细胞所得的DNA为模板进行扩增，若能得到预期长度的PCR产物，则该转化细胞就可能含有目的序列。

（3）免疫化学法　利用特定抗体与目的基因表达产物特异性结合的性质进行筛选。此法不是直接筛选目的基因，而是通过与基因表达产物的反应，指示含有目的基因的转化细胞的存在。

（4）受体/配体的结合性质　此方法也不是直接筛选目的基因，而是利用标记的配体或受体，与目的基因表达出来的蛋白质之间的相互作用来进行筛选。例如，利用过渡态类似物来筛选目的基因为酶的阳性克隆。

（5）Southwestern/Northwestern印迹法　这种方法专门用来筛选含有核酸结合蛋白基因的克隆，其中以得到DNA结合蛋白基因为目的的筛选方法称为Southwestern印迹，而以得到RNA结合蛋白基因为目的的筛选方法称为Northwestern印迹。此方法以标记的具有特定序列的双链寡聚核苷酸作为“诱饵”，筛选含有能够与此序列结合的蛋白质基因的克隆。

（6）DNA限制性内切酶图谱分析法　外源DNA插入载体会使载体DNA限制性酶酶切图谱发生变化，如果出现新的RE切点，就将转化细胞内的载体DNA抽取后酶切，进行电泳，然后观察其酶切图谱并与预期的酶切图谱相比较，从而判断转化细胞是否含有目的基因。

（7）DNA序列分析法　通过上述方法筛选得到的阳性克隆，都需要使用碱基序列测定来做最后的鉴定。已知序列的基因克隆要经序列分析确认所获得的克隆准确无误；未知序列的克隆只有在测定序列后才能了解其结构、推测其功能，以做进一步的研究。

第二节　基因克隆的详细步骤

一般基因克隆的基本步骤包括：获得外源DNA序列或目的基因，将目的基因与载体相连，将重组体导入特定的宿主细胞，筛选与鉴定含有目的基因序列的克隆。后两步请参见本章第二节。

一、外源DNA序列和目的基因的获得

无论是哪一种形式的基因克隆，第一步都是想方设法获取所需要的外源DNA序列或特定的目的基因。克隆目的不同，获取外源DNA或目的基因的手段也不同。如果克隆的目的是建立某种基因文库，就需先从目标生物细胞中抽取基因组DNA，再使用一定的方法将得到的DNA随机地切割成小的片段以后，即可进入下一步；如果是建立cDNA文库，就要从你想建立文库的细胞中抽取总mRNA，然后通过逆转录合成cDNA后进入下一步；如果是要表达一个已知的基因，首先就需要分离纯化得到该基因。

获取目的基因的手段概括起来有以下几种：①人工合成，如果一种基因较小而且碱基序列已知，就可直接使用DNA合成仪进行人工合成。如果基因较长，就需要分成几段合成，然后再连接起来。对于一级结构已知的蛋白质，可通过遗传密码表，反推出其基因的核苷酸序列，然后再进行人工合成。②使用酶切将目的基因直接从另一种克隆载体中释放出来。③逆转录，先获得某种多肽或蛋白质的mRNA，然后，通过逆转录得到以cDNA形式存在的基因。④PCR(参看本章第四节“聚合酶链式反应”)。

二、目的基因与载体的连接

将外源序列或目的基因插入载体，主要是靠 DNA 连接酶和其他工具酶的配合使用。根据末端的性质，它们的连接方式主要有 3 种：①载体和目的基因具有相同的黏端的连接；②载体和目的基因均为平端；③载体和目的基因各有一个黏端和一个平端的连接。选择哪一种连接方式主要取决于载体的性质（特别是 MCS 的性质）和目的基因的来源。

（一）黏端连接

黏端连接是最有效的连接方式。如果载体上的 MCS 含有与目的基因两端相同的 RE 切点，就可使用同一种 RE 分别消化载体和目的基因，从而在载体和目的基因上产生相同的黏端；经分离纯化后，将它们按一定的比例进行混合；低温退火后，载体和目的基因被黏端"粘"在一起；最后，在 DNA 连接酶催化下，目的基因就与载体最终以共价键相连。

有时，目的基因的两端和载体的 MCS 虽然具有不同的 RE 切点，但如果能找到能产生相同黏端的不同种 RE，同样可用此法连接。例如，识别 6 bp 序列（GGATCC）的 *Bam*H Ⅰ和识别 4 bp 序列（GATC）的 *Sau3a* Ⅰ虽然识别的序列不尽相同，但是，切割 DNA 后都产生相同的 5′突出黏端 GATC。

如果找不到合适的 RE 产生互补的黏端，就可用一些特殊的方法引入黏端。例如，可在目的基因两端添加含有特定 RE 切点的人工接头序列，也可以使用 PCR，借助事先设计好的引物，通过扩增在目的基因两端引入含有特定 RE 的序列，然后，再使用相应的 RE 消化产生黏端。

（二）平端连接

T4 DNA 连接酶可直接将含有平端的载体和含有平端的目的基因连接在一起，但连接效率要比黏端连接低得多，因此，一般很少使用这种连接方法。

（三）含有平端和黏端的目的基因与载体之间的连接

进行这种方式的连接最为少见，因为产生上述末端的可供选择的 RE 很少。但通过这种连接，目的基因只能以一种方向插入到载体之中，从而可以实现定向克隆。

第三节　基因克隆的应用

目前，基因克隆主要应用在文库建立、序列分析、表达外源蛋白、制备转基因动物和植物、基因治疗、基因敲除、基因敲减和寻找未知基因。

一、文库的建立

基因克隆中的文库是指克隆到某种载体上能够代表所有可能序列，并且可以稳定维持和使用的 DNA 片段的集合。根据序列的来源，文库可分为基因组文库和 cDNA 文库，两类文库的比较参看表 44-2。

► 表 44-2　基因组文库和 cDNA 文库的比较

文库类型	基因组文库	cDNA 文库
来源	基因组 DNA	mRNA
变化	物种	物种，组织，不同的发育阶段
插入大小	12~20 kb	0.2~6 kb
代表性	均等	与表达水平有关
类型	只有一种	两种（表达型和非表达型）
探针	DNA	DNA、蛋白质或抗体
用途	基因结构，推断蛋白质性质	表达的蛋白质，推断蛋白质性质

建立文库的主要目的在于使用合适的方法，从文库中鉴定出特定的目的序列，并进行扩增分离，此过程称为文库筛选；也可以在鸟枪序列分析（shotgun sequencing）中，随机选择一个克隆对其进行鉴定。

一个好的文库应该具备以下特点：①完整性，不遗漏任何序列；②准确性；③稳定性；④满足筛选一个重组体所需要的最低克隆数目；⑤容易筛选、贮存和扩增；⑥如果是基因组文库，就需要克隆重叠，以便通过"步移"对序列进行拼装，防止或最大限度地降低非临近序列片段连接在一起形成嵌合体克隆。

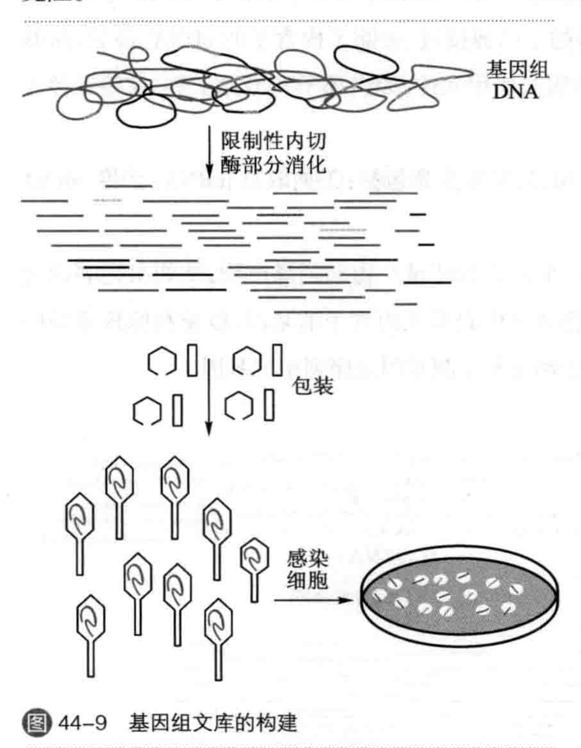


图 44-9　基因组文库的构建

（一）基因组 DNA 文库

基因组 DNA 文库可简称为基因组文库，它由一种生物的基因组 DNA 制备而来。它覆盖了一个基因组所有的序列，这些序列应该是精确无误的。

制备基因组文库的基本步骤包括（图 44-9）：①分离基因组 DNA，一般情况下，对多细胞生物而言，需要的基因组 DNA 可以从任何细胞中抽取；②插入序列的制备，使用酶法（限制性内切酶完全消化或部分消化）或物理方法（超声波处理或搅拌剪力），将基因组 DNA 切成预期的片段；③根据插入序列的大小，选择合适的载体进行克隆。选择载体与插入序列大小的大致原则是：质粒载体约 10 kb；λ 噬菌体载体为 9~23 kb；黏粒约为 40 kb；P1 噬菌体载体为 100 kb；BAC 为 100~300 kb；YAC 为 0.5~3 Mb。

一个好的基因组文库应有助于从中分离一个完整的基因或一段序列，有助于基因组序列分析，有助于了解和确定基因的组织和基因组的结构以及疾病与基因突变之间的关系，有助于对可能的基因序列、启动子、编码的蛋白质和其他性质进行预测和分析。

在建立文库的时候，可以根据一个单倍体基因组的大小除以文库中插入序列的平均大小，粗略地估计出包含一个完整的基因组序列所需要的独立克隆的数目。例如，一个以质粒为载体的人类基因组文库，插入序列的平均长度为 2 kb，单倍体人类基因组大小为 3×10^9 bp，按照上面的方法，至少含有 1.5×10^6 个独立的克隆才能够代表一个完整的人基因组序列；如果以 λ 噬菌体为载体，插入序列的平均大小为 17 kb，那代表一个完整的人基因组序列就需要 1.8×10^5 个独立的克隆；如果以 BAC 为载体，插入序列的平均大小为 200 kb，那代表一个完整的人基因组序列就需要 1.5×10^4 个独立的克隆。

如果要更准确地计算出包含一个完整基因组序列所需要的独立克隆数目，就可使用公式：$N=\ln(1-p)/\ln(1-f)$。式中 N 表示在一个基因文库中达到预期概率所需要的独立克隆数；p 代表所需片段在基因文库中出现的概率；f 代表插入到载体中序列平均大小占基因组大小的分数。

以人基因组文库为例，如果需要一个目的基因在库中出现的概率大于 99%，就以 λ 噬菌体为载体，插入序列的平均大小为 17 kb，按照上面的公式可计算出，代表一个完整的人基因组序列需要 8.1×10^5 个独立的克隆。

如果要制备一个克隆重叠的基因组文库，那需要的克隆数就是上面计算结果的 5 倍。

构建基因组文库，再用分子杂交等技术从库中钓取特定基因克隆的方法称为鸟枪法，意味着从含有众多的基因序列克隆群中去获取目的基因或序列。当基因组比较小时，此法较易成功；当基因组很大时，构建其完整的基因组文库就很难，再从庞大的文库中去克隆目的基因更是难上加难。

（二）cDNA 文库

cDNA 文库至少与基因组文库有 3 个重要的差别：首先，cDNA 文库代表的是一种单细胞生物，或者一种多细胞生物某种特定的细胞或组织内表达的所有 mRNA 的序列。由于基因组含有的奢侈基因呈组织特异性表达，且在不同环境条件、不同分化时期表达的种类和强度也不尽相同，因此 cDNA 文库具有明显的组织特异性。显然，cDNA 文库比基因组 DNA 文库小得多，能够比较容易地从中筛选出阳性克隆，并得到细胞特异表达的基因。其次，对真核细胞来说，从基因组文库中获得的基因是带有内含子和外显子的基因组基因，而从 cDNA 文库中获得的是已剪接过、去除了内含子的 cDNA。最后，从基因组文库中，还可以获得调节一个基因表达的各种顺式作用元件，如启动子和增强子等，这些元件在 cDNA 文库通常是缺乏的。

Quiz5 如果不让你使用核糖核酸酶 H，但让你使用核酸酶 S1，你如何建立 cDNA 文库？

cDNA 的合成和 cDNA 文库的构建参看图 44-10，其基本步骤包括：①抽取总 mRNA；②将 mRNA 逆转录成 cDNA；③将 cDNA 导入到特定的载体。

利用 cDNA 文库，可以进行以下工作：①确定一个基因的转录产物和翻译产物；②如果是表达文库，就可用来表达不同的蛋白质以满足各种需要；③从库中获得无内含子的基因，以便在原核系统中进行表达；④体外合成 mRNA；⑤合成探针；⑥有助于确定和预测基因组序列中的基因。

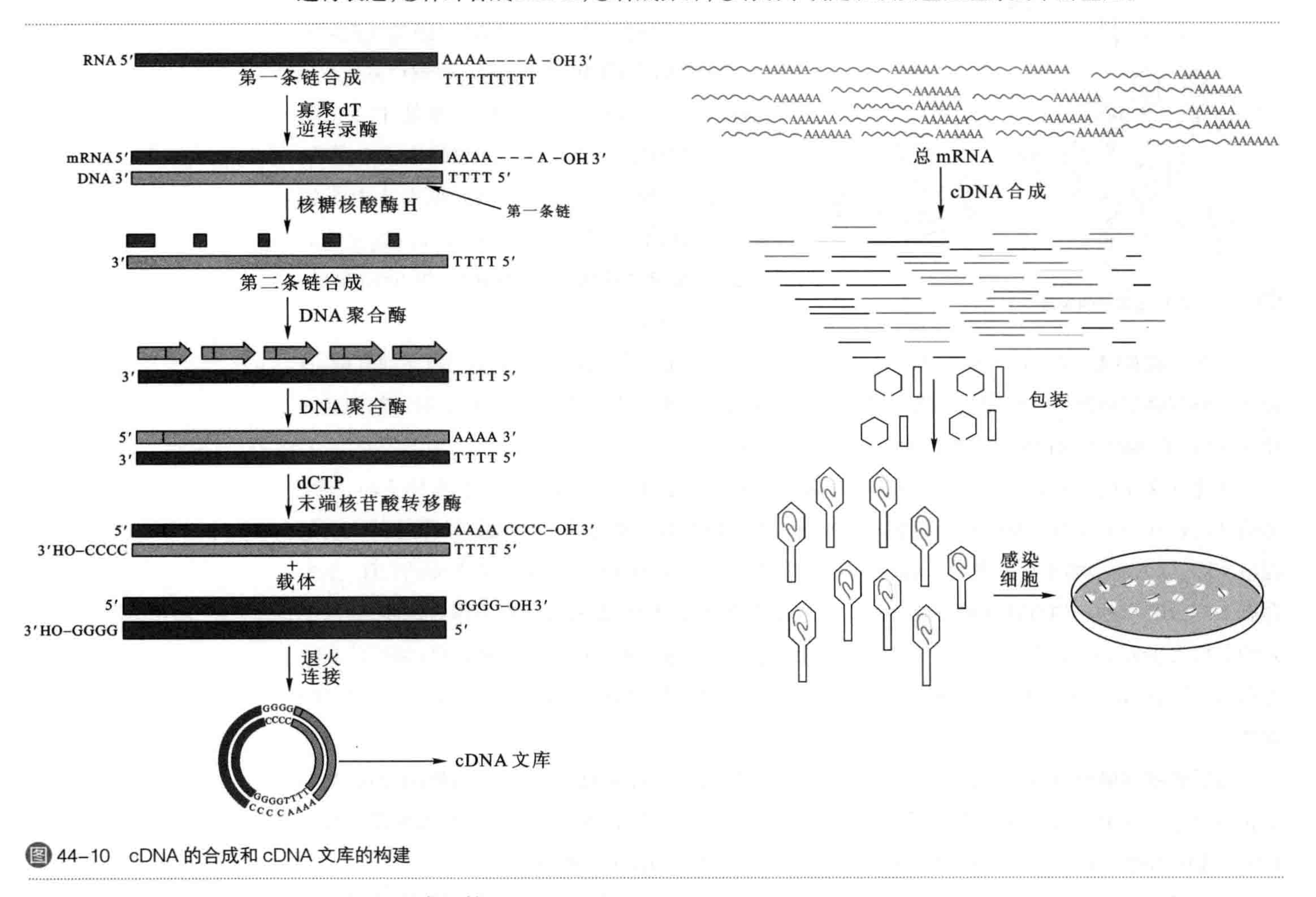


图 44-10 cDNA 的合成和 cDNA 文库的构建

（三）文库的筛选

无论是质粒文库还是噬菌体文库，文库筛选的基本原理和步骤是相同的，主要由六大步组成：①将菌落（质粒文库）或噬菌斑（噬菌体文库）复制到滤膜上；②用裂解细菌细胞壁的溶液处理滤膜，变性 DNA；③加热、烘干滤膜，以使单链 DNA 与滤膜永久性结合；④将制备好的探针与滤膜保温；⑤洗掉没有结合的探针；⑥使用放射自显影技术或其他检测系统进行最后的鉴定。

如果是表达文库，就可使用特定的抗体对表达出的蛋白质产物进行检测，也可以使用基因芯片技术对基因产物的差异表达（differential expression）进行测定；如果是非表达的基因组文库，就可使用染

色体步移(chromosome walking)法进行确定。步移法的原理是:如果一段邻近的序列已知,就可以以此段序列为起始点,分离相邻的基因,每获得一段新的序列,都可以以新得到的序列为探针,进行新一轮的筛选。

探针的来源包括:①异源探针,如果目的基因是高度保守的,就可以使用另外一个已知物种的基因序列制备探针;②cDNA 探针,由 mRNA 逆转录而成;③根据蛋白质的氨基酸序列,制备探针;④人工合成寡聚核苷酸(短单链 DNA);⑤RNA 探针,通过体外转录系统(SP6 或 T7)合成;⑥单克隆抗体。

二、DNA 序列分析

基因克隆的另一个主要目的是 DNA 序列分析。分析的对象可以是一个基因片段、一个基因、基因表达的调控序列乃至一个基因组。通过序列分析可以推测出一个蛋白质基因所编码的氨基酸序列,这有助于对一个蛋白质的性质、结构和功能进行预测;序列分析还有助于对基因和基因组的组织以及它们进化过程的理解;此外,通过序列分析,可以确定控制一个基因表达的各种顺式元件以及导致疾病发生的基因突变。

三、表达外源蛋白

使克隆的基因在特定的宿主细胞中表达,对于研究一个基因的功能及其表达调控的机理十分重要,其表达出的蛋白质可用作结构与功能的研究。许多具有特定生物活性的蛋白质或酶,具有广泛的医学或工业应用价值,将相关基因克隆之后使之在特定宿主细胞中大量表达,可满足医学或工业等领域的应用需要。

要使克隆基因在宿主细胞中表达,首先需要将目的基因亚克隆到带有基因表达所必需的各种顺式元件的载体之中,这些载体通称为表达载体(expression vector)。目的基因可以放在不同的宿主细胞中表达。针对不同的表达系统,需要构建不同的表达载体。

表达载体可分为融合载体(fusion vector)和非融合载体 2 类,前者在插入位点上“预装”了另外一个蛋白质或多肽的基因,因此,插入的外源基因将会与它发生融合,表达出来的是一种融合蛋白。例如,*lacZ* 融合序列载体、融合有蛋白 A 的 pGEX 系列(protein A series)、融合有 GFP 的 pGFP 系列、融合有多聚组氨酸标签(His-tag)的 pGEM2T 系列等。使用融合载体的主要好处是方便了目的蛋白的鉴定和纯化。

Quiz6 如果让你在大肠杆菌细胞中表达一种嗜热菌的Ⅱ型限制性内切酶,你应该如何表达才能成功?

理想的表达系统应该满足以下条件:①表达载体具有合适的 MCS,以便使外源基因能插入到正确的表达位置,或者至少是含有 3 个 ORF 的系列载体;②表达产物能够形成正确的翻译后修饰和三维结构;③为可诱导的表达系统,允许细胞生长和诱导表达,防止毒性蛋白质的积累;④表达产物易于分离和纯化;⑤表达产物最好能分泌到胞外。

Quiz7 如何能让一种蛋白质分泌到胞外?

大肠杆菌是目前应用最广泛的蛋白质表达系统。除了具备一般克隆载体所具有的基本要素以外,在大肠杆菌中使用的表达载体还应该注意“入乡随俗”,使载体带有大肠杆菌基因表达所必需的各种顺式元件,包括转录起始必需的启动子和翻译起始所必需的 SD 序列等。而外源基因最好使用的是大肠杆菌所偏爱的密码子。

然而,并不是所有的基因都适合在大肠杆菌或其他细菌中表达,当将真核基因放入原核细胞中表达时,通常面临以下问题:①缺乏真核基因转录后加工的功能,不能进行 mRNA 前体的剪接。因此,表达基因一般不能直接来源于真核生物基因组,而是来自其 cDNA。②缺乏真核生物翻译后加工的功能,导致表达产生的蛋白质,不能进行糖基化、磷酸化等修饰,难以形成正确的二硫键和空间构象,因而产生的蛋白质经常没有活性或者活性不高。③表达的蛋白质经常是不溶的,会在细菌内聚集成包涵体(inclusion body)。

使用真核表达系统表达真核生物的蛋白质,自然比原核系统优越,常用的有酵母、昆虫和哺乳动

物培养细胞等表达系统。真核表达载体至少具备2类元件:①原核生物的复制起始区和能用在细菌中筛选克隆的抗药性标记基因等。②在真核细胞中复制和表达所需要的元件,包括真核的复制起始区、启动子、增强子、转录终止和多聚A加尾信号序列、mRNA前体剪接信号序列、能用在宿主细胞中筛选的标记基因以及供外源基因插入的MCS等。

Quiz8 如何设计一个实验装置,以GFP为报告蛋白,来检测环境中雌激素的含量?

目前市场上已有多种利用DNA重组技术生产的多肽药物和疫苗销售,例如,胰岛素、干扰素、红细胞生成素(EPO)、表皮生长因子和乙型肝炎表面抗原(HBSAg)疫苗等。

四、转基因动物和植物

克隆的基因不仅可以导入细菌或培养的细胞,还能转入动植物体内,并整合到基因组中,使机体所有的细胞都带有特定的外源基因,从根本上改变了一种生物的遗传特性。转基因动物或植物就是指在其基因组内稳定地整合有外源基因、并能遗传给后代的动物或植物。

e44-3 转基因动物和植物

五、基因治疗

人类疾病的基因治疗(gene therapy)是指将人的正常基因或有治疗作用的基因通过一定方式导入到人体靶细胞,表达有功能的蛋白质,以纠正目的基因的缺陷或者发挥治疗作用,从而达到治病目的的一种治疗方法。

根据治疗的细胞对象,基因治疗分为性细胞基因治疗(germ-line gene therapy)和体细胞基因治疗(somatic gene therapy)。性细胞基因治疗是在患者的性细胞中进行操作,用来彻底根除并使其后代从此再也不会得这种遗传疾病。然而,由于目前的技术水平有限,难以解决关键基因的定点整合问题,加之相关的伦理学问题,还不能进入临床试验。体细胞基因治疗系当今基因治疗研究的主流。

根据治疗途径,基因治疗可分为体内(*in vivo*)基因治疗和回体(*ex vivo*)基因治疗。前一种途径是直接往人体组织细胞中转移基因,后一种途径需要先从病人体内获得某种细胞,进行培养,在体外完成基因转移后,将成功转移的细胞扩增培养,然后重新输入患者体内。例如,1990年美国NIH的Blase R. M.使用回体基因治疗,用正常的腺苷脱氨酶(*ADA*)基因,成功治愈两位由于ADA基因缺陷导致SCID的女孩(一个四岁,一个九岁)。目前这两个女孩早已经长大,一直过着正常人的生活。

无论是体内还是体外基因转移,都需要一种安全、高效和无毒的载体将外源基因带入病变细胞,从而进入细胞核与染色体DNA整合,并获得表达。目前使用的载体有2类:一类是由腺病毒改造而成,主要用于体内基因治疗;另一类是由逆转录病毒改造而成,主要用于回体基因治疗。

六、基因敲除

e44-4 基因敲除

基因敲除(gene knockout)是指在分子水平上,使用特定的手段,将一个结构已知但功能不详的基因去除,或用其他顺序相近的基因取代,使原基因功能丧失,然后从整体观察实验生物的表型变化,进而推断相应基因的功能。这与早期生理学研究中常用的切除部分-观察整体-推测功能的思路相似。

Quiz9 当今我国最重要的模式动物(小鼠)敲除中心在哪里?

目前,基因敲除的应用领域主要有:①建立人类疾病的转基因动物模型,为医学研究提供材料。②改造动物基因型,鉴定新基因和/或其新功能。

七、基因敲减

基因敲减(gene knockdown)也称为基因抑制,是一项降低或者抑制一种生物的某个或者某些基因表达的技术,以区别传统的“基因敲除”。敲减的手段可以在DNA水平上,通过对DNA的修饰来抑制基因的转录,也可以使用人工设计的核酶(主要是锤头核酶),定向切割特定的目标基因转录出来的mRNA,或者在翻译水平上,通过RNAi技术或者依赖于核糖核酸酶H的反义核酸技术,来抑制特定mRNA的翻译。如果是在DNA水平上进行基因敲减,就可以将其与转基因技术结合起来,得到转基

因敲减生物。

传统的基因敲除的方法是让一个 DNA 片段从染色体完全剔除，或者让一个基因在某个生物体内完全灭活，这等于是让这个生物完全丧失了原来的基因。而与基因敲除不同的是，基因敲减保留了原来的基因，只是抑制了原来基因的表达。但无论是基因敲除，还是基因敲减，都会影响到基因的表达，并最终导致被敲除的或者敲减的生物在表型上发生改变，借此可以用来研究一个基因的功能。

现在用得最多的是基于 RNAi 建立起来的基因敲减。这种敲减策略的一个优势就是它几乎适用于所有动植物。现在经过改进的基于 RNAi 的敲减策略，可使特异组织中的基因表达沉默，而且可以设计特定的 RNAi，在生物的发育期或成年期的任何时间打开或关闭。

Quiz10 在图 44-11 所示的人造生命的流程中，有一步是将人造的基因组 DNA 在体外进行甲基化，这一步的用途是什么？如果是人造酵母，还需要这一步反应吗？

八、寻找未知基因

基因克隆不仅可以使人类对已知的基因进行各式各样的研究，还为我们寻找和鉴定新基因提供了一种十分有效的途径。寻找新基因的途径可以从基因的终产物着手，也可以直接在核酸水平上进行。

(一) 从基因的终产物开始鉴定新基因

以一个未知的蛋白质基因为例，如果先分离、纯化到这种新的蛋白质或者它的降解片段，就可确定它的部分氨基酸序列，然后根据遗传密码子表，反推出编码它的核苷酸序列并进行人工合成，作为探针从 cDNA 文库或基因库中挑出原始的基因。

Quiz11 为什么在根据氨基酸序列设计核酸探针序列的时候，一般集中在富含 Met 和 Trp 的区域，而忽略富含 Leu 和 Arg 的区域？

(二) 从核酸水平上寻找新基因

随着基因组学的兴起和发展，人们得到各种生物基因组的全部序列和部分序列，从这些已知 DNA 序列中得到新基因成为研究人员的一大目标。但是，由于基因结构的复杂性和多样性，很难建立通用的捷径或法则来鉴别新基因，总结起来主要有 9 种方法。

e44-5 从核酸水平上寻找新基因

框 44-1　生化新突破——人造生命的诞生

2010 年 5 月 20 日，美国科学家 Craig Venter 向世人宣布，由他领导的研究小组历时十余年，耗资超过 4 000 万美元，获得世界首例人造生命——完全由人造基因组控制的支原体，并将它命名为“辛西娅”(Synthia)。这项具有里程碑意义的实验表明，新的生命体可以在实验室里“被创造”，而不一定要通过自然进化的方式来完成。

他们获得人造生命的基本过程是(图 44-11)：首先，在体外人工合成一种名为蕈状支原体基因组

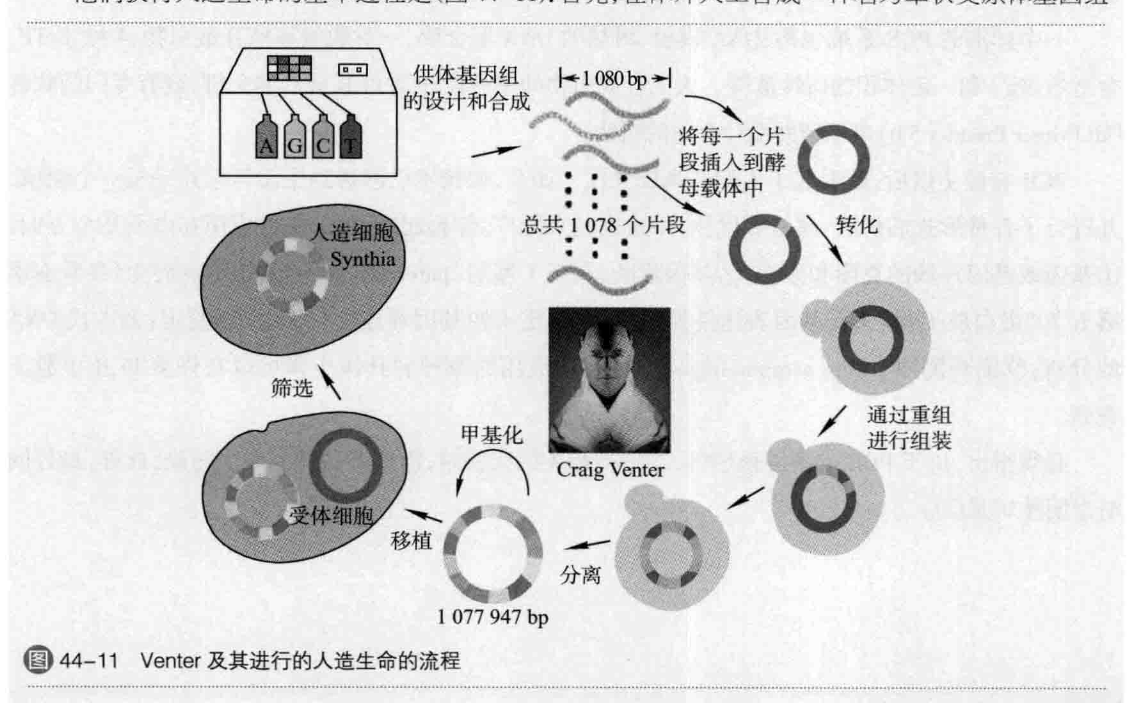


图 44-11　Venter 及其进行的人造生命的流程

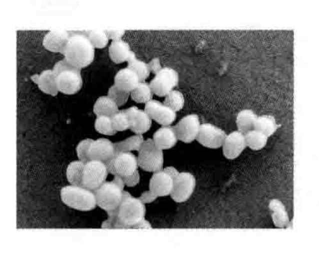


图 44-12 显微镜下的 Synthia

DNA；然后，将人造基因组移植到山羊支原体中。通过生长和分裂，受体支原体将一分为二，产生2个细胞，一个带有人造的基因组 DNA，另一个带有天然的基因组 DNA；通过培养皿中的抗生素，将带有天然基因组 DNA 的细胞杀死，只留下人造的支原体细胞不断增殖，于是新的生命便诞生了（图 44-12）。

Venter 还事先在人造基因组 DNA 上写入了 4 个独特的“水印序列”，使其有别于同类的天然支原体，以及能在这种生物的后代中识别它的“祖先”。

面对实验的成功，Venter 曾情不自禁地说，“当带着水印的细胞活了过来，我们欣喜若狂，它是一个活生生的生物了，成为了我们地球上各种生命的一部分。”

虽然这种技术目前仍处于实验阶段，但研究人员相信其运用前景广阔。辛西娅的基因组中仅包含了约 850 个基因，这些基因都是维持一个独立的生命存在所必需的。在此基础上，Venter 打算将来可以通过向其添加其他基因，制造一系列新的微生物，让它们可生产生物燃料、有用的药品，可以从空气中吸收二氧化碳和其他污染物，或者制造合成疫苗所需的蛋白质。

第四节 聚合酶链式反应

聚合酶链式反应（polymerase chain reaction，PCR）是一种在体外特异性扩增特定 DNA 序列或片段的方法，由美国科学家 Kary Mullis 于 1984 年发明。

PCR 的原理并不复杂：理论上，DNA 分子数目经复制呈指数增长，如果提供足够的引物和 dNTP，1 分子 DNA 复制 n 次后，就可产生 2^n 个 DNA 分子。但与体内 DNA 复制不一样的是：PCR 的解链反应使用的是热变性，而不是解链酶；PCR 使用的引物是人工合成寡聚 DNA，而不是体内由引发酶合成的 RNA；为了增加 DNA 聚合酶的稳定性，PCR 使用的是耐热的 DNA 聚合酶。

整个 PCR 反应由多个循环组成（循环次数为 30 ~40 次），每循环一次，DNA 复制一次。每一个循环由 3 步反应组成（图 44-13）：① DNA 变性——采取热变性，使模板 DNA 在 95℃左右的高温下解链；②退火——降低温度（通常在 50~65℃），使引物与模板 DNA 配对；③延伸反应——在 DNA 聚合酶催化下，在引物的 3′端合成 DNA，温度通常在 72℃左右。在循环结束以后，一般还有一步专门的延伸反应，持续 10~30 min。最后得到的 PCR 产物可以通过常规的琼脂糖凝胶电泳进行鉴定分析。

Quiz12 PCR 最后一步是 10~30 min 专门的延伸反应，其目的是什么？

一个标准的 PCR 系统包括：DNA 模板、耐热的 DNA 聚合酶、一对脱氧寡核苷酸引物、4 种 dNTP、合适的 Mg^{2+} 和一定体积的缓冲液等。人工合成引物的序列设计是 PCR 成功的关键，现有专门的软件（如 Primer Premier 5.0）可以辅助设计合适的引物。

PCR 自诞生以后，就引起了人们的高度关注。如今，该技术已渗透到生命科学几乎每一个领域，并进行了各种形式的扩展、改进和优化，用途也越来越广。综合起来，PCR 主要应用在以下几个方面：①基因或基因片段的克隆和鉴定；②基因诊断；③亲子鉴定（paternity testing）；④定点突变（参看本章第五节“蛋白质工程”）；⑤基因表达差异定量；⑥确定未知基因表达变化；⑦法医鉴定；⑧古代 DNA 的分析；⑨循环测序（cycle sequencing）。上述各项应用的原理和具体步骤可以在许多 PCR 手册上查到。

e44-6 各种 PCR 技术的原理及其运用

必须指出，由于 PCR 的高度敏感性，在进行相关的实验时，需要严防样品发生污染，此外，最好同时做阴性对照反应。

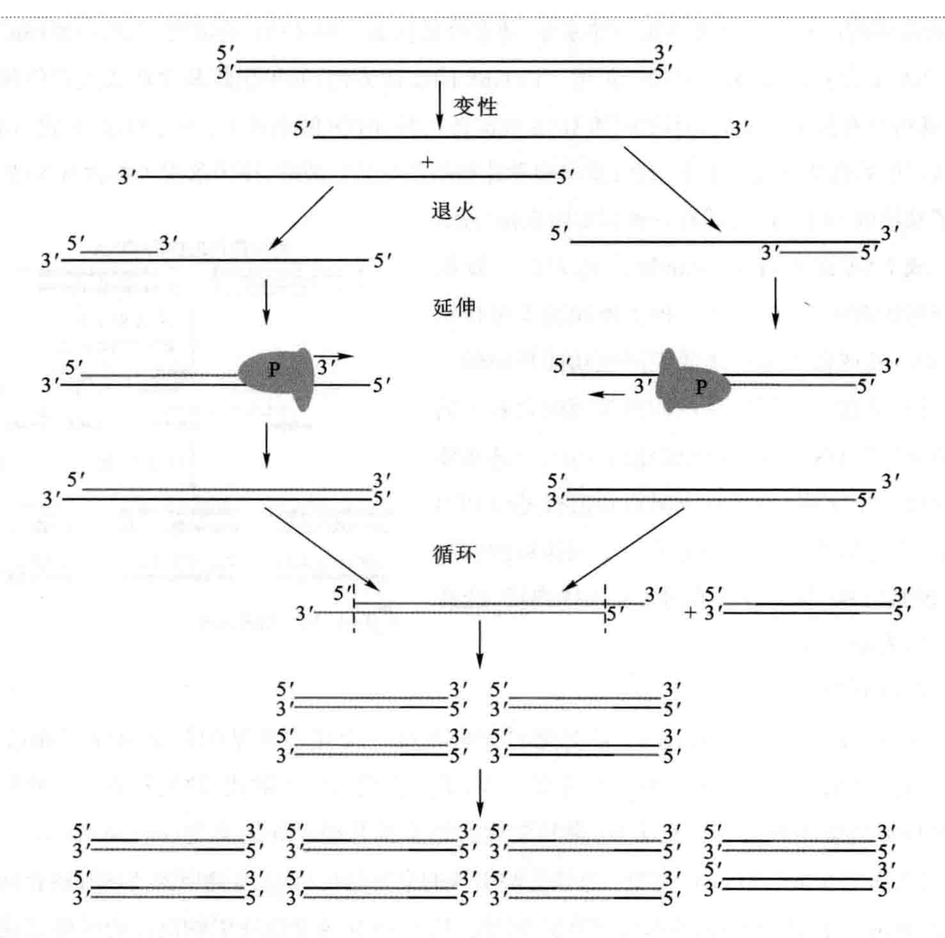


图 44-13 聚合酶链式反应的基本过程

第五节 蛋白质工程

蛋白质工程(protein engineering)又称为第二代基因工程,是使用遗传学和化学手段,从改变或合成基因入手,改善一种蛋白质的结构与功能,从而产生具有特殊功能、符合人们意愿性质的新产物的一项技术。

蛋白质工程一般有4个目的:①改变或改良催化性质。②改变结构性质。③创造新系统,包括合成融合蛋白或多功能蛋白、添加有利于纯化的标签或增强药用蛋白的药效等。④蛋白质的定向进化(directed evolution),在不需要事先了解蛋白质的三维结构和作用机制的情况下,直接在体外模拟自然进化的过程(随机突变、重组和选择),使基因发生大量变异,并定向选择出所需性质或功能的蛋白质,在较短时间内完成漫长的自然进化过程。

改造蛋白质的主要手段是体外突变(*in vitro* mutagenesis)。突变又分为非特异性的随机突变和特异性的定点突变,只有特异性的定点突变才是按照人们的事先设计进行的,具有明确的目的,因此成为蛋白质工程的主要手段。

(一) 随机突变

随机突变可以通过合成带有随机取代的寡核苷酸片段,进行寡核苷酸引导的突变(参考本节"寡核苷酸引导的定点突变")。也可以使用化学诱变剂直接作用DNA,制作突变序列文库,然后进行筛选。但使用最方便的还是PCR突变。

PCR突变可以是随机的,也可以是定向的。若是随机突变,可以在特定的条件下进行易错的PCR(error-prone PCR)。例如,在扩增体系中用Mn^{2+}取代Mg^{2+},因为在DNA复制的时候,使用Mn^{2+}代替

Mg^{2+}，可提高错配的机会。或者在反应体系中，故意降低任意一种 dNTP 的浓度，也可以使用缺乏校对活性的 DNA 聚合酶在高 Mg^{2+} 下进行扩增。以 GFP 的改造为例，如果想将其改造成发其他颜色荧光的 XFP，就将含有野生型 *gfp* 基因的载体 DNA 在缺乏一种 dNTP 的条件下，进行 PCR，并进行克隆、转化和表达。最后在紫外线照射下，直接挑出能发其他颜色荧光的菌落，该菌落很可能含有 XFP。

除了易错的 PCR 以外，还有一种叫基因混排（gene shuffling）或 DNA 混排（DNA shuffling）的方法。基因混排也叫基因改组，实际上是一种体外同源重组技术（图 44-14）。具体操作是将来源不同但功能相同的一组同源基因（来源于不同物种的同源基因或含有不同突变的基因），用 DNA 酶 I 消化成随机片段，由这些随机片段组成一个文库，使之互为引物和模板进行 PCR 扩增，当一个基因拷贝片段作为另一基因拷贝的引物时，引起模板互换，重组从而发生，导入体内后，选择正突变体作为新一轮的体外重组。

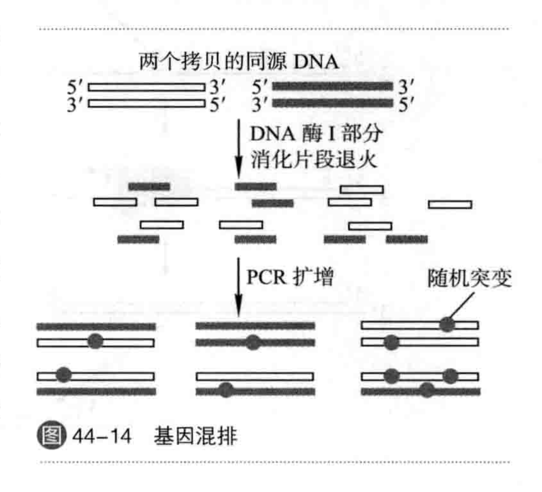


图 44-14　基因混排

（二）定点突变

定点突变（site-directed mutagenesis）是通过定向改变一个蛋白质基因的碱基序列而改变多肽链上一个或几个氨基酸的序列。与天然突变一样，定点突变也分为取代、缺失和插入 3 种形式。目前已有多种方法被用来进行定点突变，最早使用的是寡核苷酸引导的突变（oligonucleotide-directed mutagenesis），而现在多使用 PCR 突变。寡核苷酸引导的突变是在 DNA 复制所需要的寡核苷酸引物中引入突变，再通过 DNA 体外和体内复制将突变固定下来；PCR 突变也在引物设计的时候直接引入突变，但完全通过体外扩增将突变固定到一个完整的基因之中（图 44-15）。

e44-7 寡核苷酸引导的定点突变

无论是随机突变，还是定点突变，最后都面临一个筛选的问题。如何提高筛选的效率，以在较短的时间内获得想要的变体，这对于蛋白质工程的成败至关重要。现在有 2 种非常有用的筛选方法，一个是噬菌体展示，另一个就是核糖体展示（ribosome display）。

Quiz13 你如何利用噬菌体展示技术，获得对受体具有更高亲和力的人生长激素？

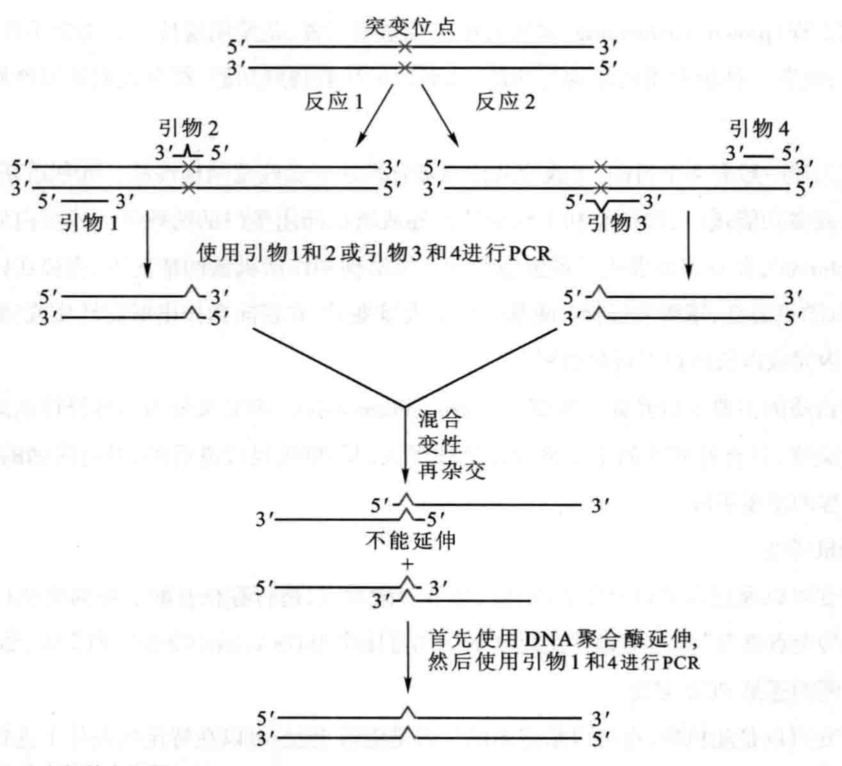


图 44-15　PCR 突变的基本流程

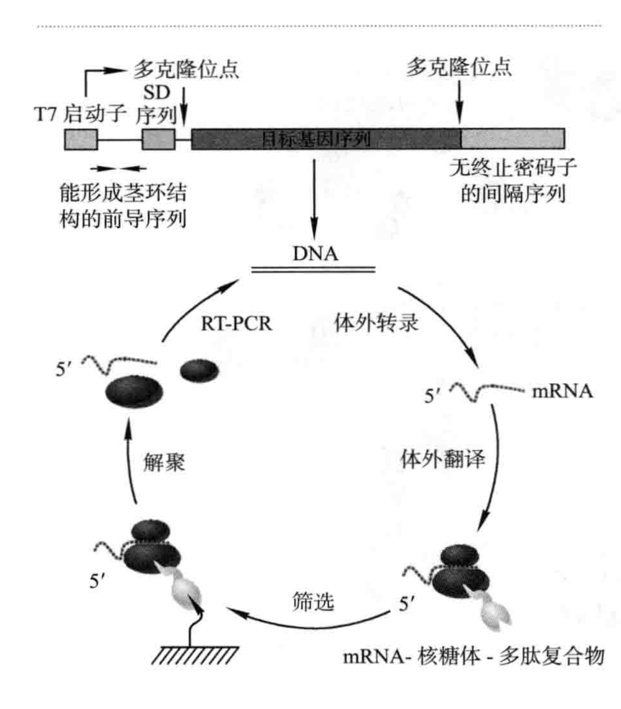


图 44-16　核糖体展示

核糖体展示是一种利用蛋白质与特定配体特异性结合而在体外直接进行筛选的技术。它将正确折叠的蛋白质及其 mRNA 模板同时结合在核糖体上，形成 mRNA- 核糖体 - 蛋白质三元复合物，从中筛选出目标蛋白和编码它的基因序列，并可用于各种蛋白质的体外改造等。

核糖体展示技术的基本原理和步骤包括(图 44-16)：①模板的构建。模板的构建需要将目标 DNA 序列(突变的或没有突变的)插入到一种受 T7/SP6 RNA 聚合酶驱动的体外转录载体之中。在构建这种载体的时候，需要将目标 DNA 的 5′端与一段能形成茎环结构的前导序列融合在一起，以提高将来转录物的稳定性。而在 3′端与一段无终止密码子的间隔序列融合在一起，使转录物缺乏终止密码子。②体外转录和翻译。在模板构建好以后，先进行体外转录，再使用无细胞翻译系统进行体外翻译。由于转录物缺乏终止密码子，故在进行体外翻译的时候无法终止。因而翻译到最后，肽酰 tRNA 仍然与间隔序列结合，而由原来的基因序列翻译出来的蛋白质突出在核糖体之外并进行折叠，由此形成一种由 mRNA、核糖体和蛋白质组成的三元复合物。③亲和筛选。根据突变的蛋白质与特殊配体特异性结合的性质，使用亲和层析将核糖体上还没有释放出来的蛋白质，与结合在特殊树脂表面的配体或包被有特殊配体的磁珠保温结合，来筛选目标蛋白。为了稳定形成的 mRNA- 核糖体 - 蛋白质复合物，可将温度降低并加入 Mg^{2+}。随后可使用高盐溶液或 EDTA 或游离的配体进行洗脱。④逆转录 PCR。mRNA 在洗脱的时候将得以释放，随后可以作为模板，进行逆转录 PCR，并进行下一轮突变、转录、翻译和筛选，以得到更好的突变体。

第六节　研究核酸与蛋白质相互作用的主要方法和技术

分子生物学的核心内容之一是研究核酸与蛋白质之间的相互作用。用来研究这两大生物分子相互作用的主要方法和技术有：电泳泳动变化分析、DNA 亲和层析、DNA 酶 I - 足印分析(参看第三十六章“DNA 转录”)和染色质免疫沉淀技术(chromatin immunoprecipitation，ChIP)。

ChIP 是研究体内蛋白质与 DNA 相互作用的有力工具，利用该技术不仅可以检测细胞内各种反式作用因子与 DNA 分子上各种顺式作用元件之间的动态作用，还可以用来研究组蛋白的各种共价修饰以及转录因子与基因表达的关系。

ChIP 需要在活细胞状态下使用甲醛固定蛋白质 -DNA 复合物，并通过超声或酶处理，将染色质随机切成一定长度范围内的小片段。然后通过抗原抗体的特异性识别和结合反应沉淀此复合体，特异性地富集与靶蛋白结合的 DNA 片段。最后通过对目的片段的纯化与检测，获得蛋白质与 DNA 相互作用的信息。

ChIP 操作的基本步骤是(图 44-17)：①用甲醛在体内将 DNA 结合蛋白与 DNA 交联；②分离染色质，使用超声法或者酶法将染色质剪切成小的片段；③先用特异性抗体与 DNA 结合蛋白结合，再用沉淀法分离形成的复合体；④解交联，纯化富集释放出来的 DNA 片段；⑤用 PCR 扩增释放出来的 DNA

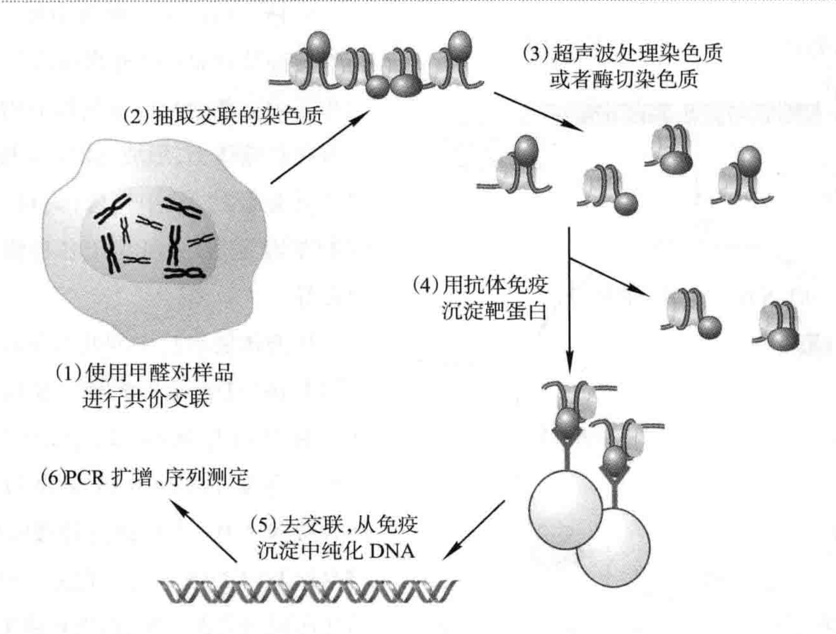


图 44-17 染色质免疫沉淀流程示意图

片段并进行序列分析。

第七节 研究蛋白质之间相互作用的主要方法和技术

分子生物学的另一核心内容是研究蛋白质与蛋白质之间的相互作用。用来研究蛋白质之间相互作用的主要方法和技术有免疫共沉淀、亲和层析、共价交联、荧光共振能量转移(fluorescence resonance energy transfer，FRET)、生物发光共振能量转移(bioluminescence resonance energy transfer，BRET)、酵母双杂交和蛋白质芯片技术(见本章第九节)。

一、免疫共沉淀

此方法的原理是：如果蛋白 X 与蛋白 Y 之间存在相互作用，那当将蛋白质 X 的抗体加到细胞裂解物之后，X- 抗体复合物就会让 Y 一并发生免疫沉淀。

二、亲和层析

此方法的原理是：如果蛋白 X 与蛋白 Y 之间有相互作用，那在将细胞裂解液流过固定有 X 的树脂以后，Y 就通过与 X 之间的特异性相互作用而被亲和吸附在树脂上，其他无关的蛋白质会直接流出树脂。

三、共价交联

此方法的原理是：如果蛋白 X 与蛋白 Y 之间有相互作用，那在细胞或细胞裂解液中加入交联试剂以后，它们之间就会形成稳定的共价复合物。然后，再使用免疫沉淀的方法将它们共同沉淀下来。最后，将交联打开，Y 得以释放。

四、荧光共振能量转移

此方法的原理是：如果两种蛋白之间有相互作用，那在将它们各自引入激发的荧光供体基团和荧光受体基团以后，两个荧光基团之间就会发生能量转移。FRET 在两个荧光基因之间的距离小于 10 nm 时就可以发生，这可以通过荧光受体发出的荧光波长的变化来测定。

五、生物发光共振能量转移

此方法的原理类似于FRET，不同的是要借助重组DNA技术将两种蛋白质分别与不同的荧光蛋白融合在一起，然后再测定它们之间的能量转移。

六、酵母双杂交

酵母双杂交是一项专门用来研究蛋白质与蛋白质相互作用的技术。它首先由S. Fields和O. K. Song于1989年提出。这项技术在验证已知蛋白质之间的相互作用或筛选与特定靶蛋白呈特异性作用的候选蛋白的研究中，已经得到了广泛的应用。其可行性和有效性已被证实，并被推广到了诸如信号转导、细胞周期调控、肿瘤基因表达等多个研究领域。

(一) 酵母双杂交系统的原理

酵母双杂交系统技术的基本原理是，激活基因表达的激活蛋白具有2种功能不同的结构域：一种是与DNA结合的结合结构域(BD)，另一种是激活DNA转录的激活结构域(AD)。这两种结构域并不一定需要在同一个蛋白质分子上才起作用。事实上，一个含有BD的蛋白质如果能够与另一个含有AD的蛋白质结合在一起，就可以激活转录(图44-18)。

在双杂交系统中，需要表达2种融合蛋白：一种是蛋白质X，用它作为“诱饵”，去捕获与它相互作用的目标蛋白，因此经常称为诱饵蛋白(bait protein)。X在N端与BD融合在一起；另一种是潜在的能够与X结合的候选目标蛋白Y。Y与AD融合在一起。如果X与Y相互作用，形成的XY复合物在功能上就相当于一个完整的单一激活蛋白，就能够驱动一个容易检测的报告基因(如GFP和β-半乳糖苷酶的基因)的表达。于是，报告基因的表达量可以用来作为测定X与Y相互作用的尺度。

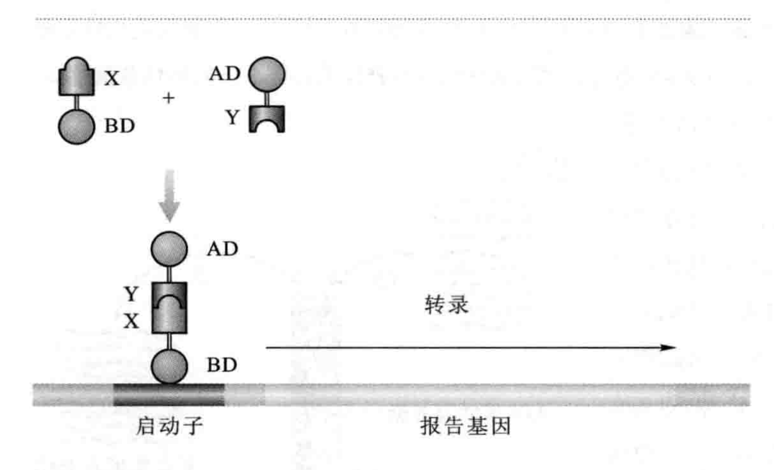


图44-18 酵母双杂交系统的原理图解

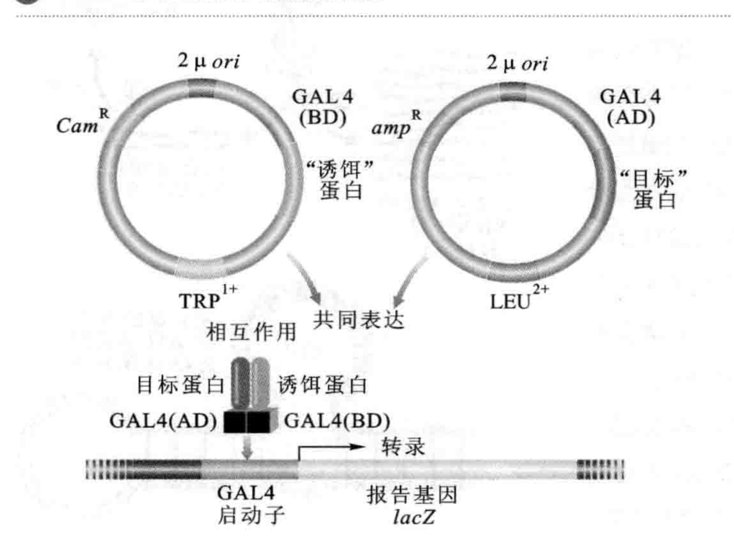


图44-19 基于GAL4的BD和AD构建的双杂交系统(图中的2μ ori代表酵母的2μm质粒复制起始区)

(二) 酵母双杂交系统的建立

双杂交系统建立的基本步骤与普通的克隆实验差不多，可分为以下几步：

(1) 选择载体 进行双杂交筛选，首先需要选择合适的载体，目前已有各种商业化的含有BD或AD的载体可供使用。而使用最多的载体上的BD或AD来自酵母的GAL4蛋白，也有一些载体上的BD来自大肠杆菌的LexA蛋白，AD来自单纯疱疹病毒的激活蛋白VP16。无论是BD载体，还是AD载体，它们都含有合成特定营养成分(通常是氨基酸)所需要的某一种酶的基因，以提供选择性标记(图44-19)。

(2) 构建融合基因 将“诱饵”蛋白X的基因和目标蛋白Y的基因分别插入到BD载体和AD载体之中，以形成BD-X和

AD-Y 融合基因。

(3) 转染 使用特定的手段将重组后的 BD-X 载体和 AD-Y 载体转染到特定的营养缺陷型酵母宿主细胞。

(4) 筛选 利用双营养缺陷型的恢复筛选出同时含有 BD 载体和 AD 载体的细胞。

Quiz14 你如何利用酵母双杂交技术，筛选能够加强或减弱细胞内两种蛋白质相互作用的药物?

(5) 活性检测 一旦 BD-X 载体和 AD-Y 载体进入宿主细胞，如果 X 蛋白和 Y 蛋白发生相互作用，宿主细胞内的报告基因就可能被驱动表达。使用 GAL4 系统的报告基因通常是 *lacZ*，一旦它被驱动表达，则细胞在含有 X-gal 的培养基上显色为蓝色。

第八节 SELEX 技术

SELEX 是指指数富集式配体系统进化(systematic evolution of ligands by exponential enrichment, SELEX)。它是一项将寡核苷酸扩增和体外筛选结合在一起的技术，已被广泛用于分离能与靶蛋白或小分子高亲和力结合的 RNA 和 DNA 分子。筛选到的寡核苷酸序列称为适体。能与靶蛋白分子作用的适体可用硝酸纤维素滤膜捕获，或者以初始聚丙烯酰胺凝胶分离，能与小分子结合的适体则可用亲和层析分离。该技术不仅可用于研制核酸药物，还可用于研制新的核酶。例如，利用 SELEX 技术筛选获得的针对血管内皮生长因子(VEGF)的一种 RNA 适体，经过美国 Eyetech 和 Pifzer 制药公司 15 年的研制，已经于 2004 年 12 月获得美国 FDA 批准上市，用于治疗老年性视网膜黄斑营养不良。

SELEX 筛选过程是应用大容量的随机寡核苷酸文库，结合 PCR 进行体外扩增，以指数级富集与特定靶分子特异结合的寡核苷酸，经过数轮反复的体外筛选、扩增，最终获得与靶分子特异结合的寡核苷酸适体。核酸文库由组合化学合成制备，主要有 RNA 文库、DNA 文库和含有修饰核苷酸的文库。文库中由中间为一定长度的随机序列和 5′端、3′端的固定序列构成。两端固定序列长度一般为 20~25 nt，其作用是增加文库的稳定性和为 PCR 扩增准备。随机区的每一个核苷酸位置都存在 4 种可能性，如果随机序列长度为 N，随机序列的多样性就有 4^N 种，即文库的库容量为 4^N。典型的文库至少有 10^{13}~10^{18} 种独立序列，其随机区域长度为 30 nt 左右。

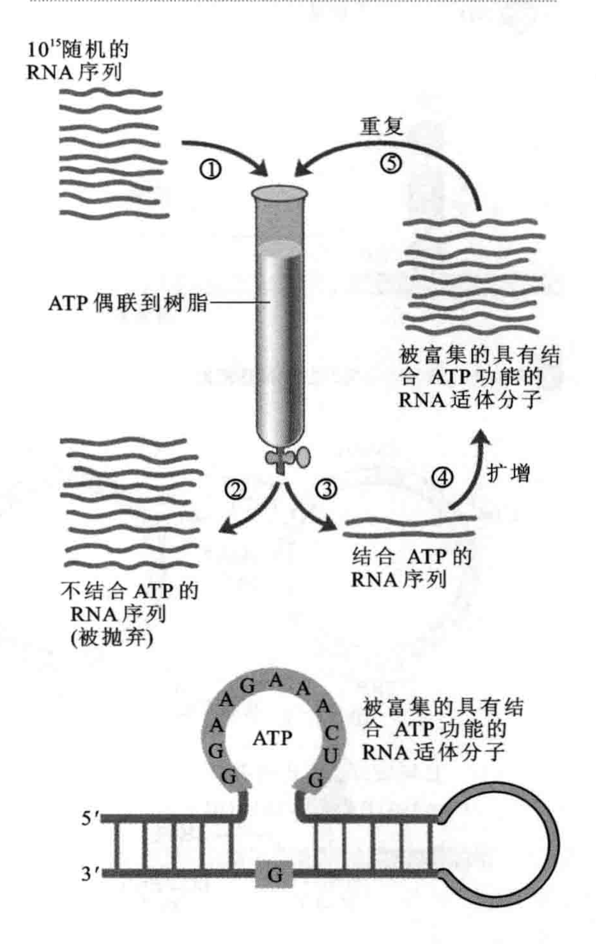


图 44-20 SELEX 技术筛选 ATP 结合 RNA 流程(Nelson 等,2008)

筛选流程是将文库和靶分子在一定温度下(通常为 37℃)保温。在最初几个循环中，只有少数(0.1%~0.5%)的序列与靶分子作用，这些序列通过亲和层析、纤维膜过滤等分离手段将结合复合物与未结合的序列分开。分离得到的序列再通过 PCR 扩增产生次级文库，用于下一轮筛选，由此进入一个反复筛选富集的过程。对于 RNA 文库，扩增步骤还包括：RNA 序列的反转录以得到双链 DNA 文库，再转录生成次级 RNA 文库。经过数轮的筛选富集，如果分离得到的随机序列与靶物质亲和力不再提高，就可对筛选获得的寡核苷酸适配子进行克隆、测序、生物活性及分子识别等功能研究。

以制备结合 ATP 的 RNA 适体为例(图 44-20):

首先合成大约有 10^{15} 种随机序列的 RNA 多聚物混合物；然后将混合物通过结合有 ATP 的树脂进行选择，丢弃直接流出树脂的 RNA；对与树脂结合的 RNA 进行洗脱和收集；收集的 RNA 通过反转录 PCR，扩增得到 cDNA；将 cDNA 再转录大量的 RNA；将得到的 RNA 再进行下一轮选择和富集，这可进行多次循环，直至得到少数能够与 ATP 紧密结合的适体分子。

第九节　生物芯片技术

生物芯片是指包被在硅片、尼龙膜等固相支持物上的高密度的核酸、蛋白质、糖类、细胞、组织或其他生物组分的微阵列。将芯片与标记的样品进行杂交，通过检测杂交信号即可实现对生物样品的分析。由于常用硅芯片作为固相支持物，且在制备过程运用了计算机芯片的制备技术，因此称之为生物芯片技术。目前常见的生物芯片主要有基因芯片(gene chip)、蛋白质芯片(protein chip)和组织芯片。其中基因芯片和蛋白质芯片应用相对广泛。

一、基因芯片

基因芯片是随着"人类基因组计划"和其他模式生物基因组计划的进展而发展起来的一项技术，也叫 DNA 芯片、DNA 微阵列(DNA microarray)或寡核苷酸阵列(oligonucleotide array)。它采用原位合成(*in situ* synthesis)或显微打印手段，将数以万计的 DNA 探针固定在支持物的表面，产生二维 DNA 探针阵列。然后，将其与标记的样品分子进行杂交，通过检测杂交信号的强弱，对样品进行快速、并行和高效的检测或医学诊断。基因芯片以其无可比拟的信息量、高通量、快速和准确地分析基因的本领，在基因组功能研究、临床诊断及新药开发等方面显示出巨大的威力，已成为人类研究生命和维护生命的一大利器。

一般说来应用基因芯片分 2 步进行：①生物学问题的提出和芯片设计与制备；②样品制备；③核酸杂交反应；④结果探测；⑤数据处理和建模。

1. 基因芯片的种类和制备

目前市场上基因芯片制备的方法有 2 大类：一类是原位合成的基因芯片，另一类为直接点样而制成的基因芯片。前者适用于寡核苷酸，后者多用于大片段 DNA，有时也用于寡核苷酸，甚至 mRNA。原位合成有光蚀刻法和喷印法。点样法较简单，只需将预先制备好的寡核苷酸或 cDNA 等样品，通过自动点样装置点于经特殊处理的玻璃片或其他材料上即可。

2. 基因芯片样品的制备

在待分析基因与基因芯片结合杂交之前，需要对样品进行分离、扩增和标记。采用的基因分离、扩增及标记方法因样品来源、基因含量及检测方法和分析目的不同而不同。虽然常规的基因分离、扩增及标记技术仍旧可以使用，但现已逐步被专门的高度集成的微型样品处理系统取代。而为了获得基因的杂交信号，必须对目的基因进行标记。目前广泛使用的标记方法是荧光标记，可以采用的荧光染料种类很多，一般能够满足对不同来源的样品进行平行分析的需要。使用计算机控制的高分辨荧光扫描仪可获得结合于基因芯片上的目的基因的荧光信号，通过计算机处理即可给出目的基因的结构或表达信息。

3. 基因芯片杂交

在样品制备好以后，就可以进行下一步杂交反应了。杂交条件的选择与研究目的有关：如果芯片用于多态性分析或者基因测序，就需要将每个核苷酸或突变位点都检测出来。这时通常设计出一套 4 种寡聚核苷酸，在目标序列上跨越每个位点，只是在中央位点碱基有所不同，根据每套探针在某一特点位点的杂交严紧程度，即可测定出该碱基的种类。由于突变检测要鉴别出单碱基错配，因此需要更

高的杂交严紧性和更短的时间。如果芯片仅用于检测基因表达，就只需设计出针对基因中的特定区域的几套寡聚核苷酸即可。

此外，杂交反应还必须考虑反应体系中的盐浓度、探针的GC含量和所带电荷、探针与芯片之间连接臂的长度及种类、待测基因的二级结构的影响等因素。

4. 基因芯片检测原理

杂交信号的检测是基因芯片技术中的重要组成部分。由于基因芯片本身的结构和性质的特殊性，需要确定杂交信号在芯片上的位置，尤其是大规模基因芯片由于其面积小、密度大、点样量又很少，因此杂交信号较弱，需要使用光电倍增管或冷却的电荷偶连照相机（CCD）或摄像机等弱光信号探测装置。此外，大多数DNA芯片杂交信号谱型除了分布位点以外还需要确定每一点上的信号强度，以确定是完全杂交还是部分杂交，因而探测方法的灵敏度及线性响应也是非常重要的。

5. 基因芯片的应用

(1) 基因表达分析　基因芯片具有高度的敏感性和特异性，它可以监测细胞中几个至几千个mRNA拷贝的转录情况，可自动、快速地检测出成千上万个基因的表达情况。它不仅可以检测和分析基因表达时空特征、基因差异表达，还可用来发现新基因。

进行基因表达分析的基本步骤包括(图44-21)：① RNA的抽取和分离。先得到总mRNA，然后使用寡聚dT作为引物，在逆转录酶催化下得到cDNA。②扩增。使用T7 RNA聚合酶和生物素标记的UTP和CTP，体外转录cDNA，得到大量的生物素标记的互补的RNA（cRNA）。③“碎片化”。将cRNA保温在94℃的缓冲溶液中，产生35~200 nt长的cRNA片段。④杂交。将芯片与cRNA杂交，随后洗去非杂交的原料。⑤染色和洗脱。使用链霉亲和素－藻红蛋白对生物素标记的cRNA进行标记，然后，洗去非特异性结合的染料。⑥使用共聚焦激光扫描装置扫描杂交芯片。⑦信号放大。使用山羊抗体和生物素标记的抗体与芯片保温，再进行染色和洗脱。⑧再次扫描芯片，并对表达状况进行定量分析。

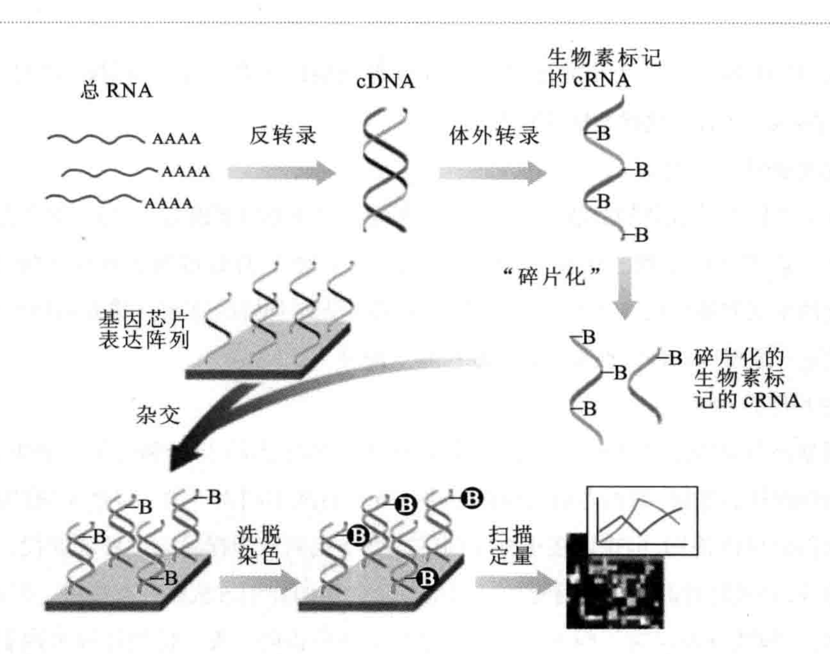


图44-21　使用基因芯片进行基因表达分析的基本流程

(2) 基因型、基因突变和多态性分析　在同一物种不同种群和个体之间，有着多种不同的基因型，而这种不同往往与个体的不同性状和多种遗传疾病有着密切的关系。通过对大量具有不同性状的个体的基因型进行比较，就可以得出基因与性状的关系。但是，由于大多数性状和遗传性疾病由多个基因同时决定，因此分析起来就很困难，然而基因芯片技术却能很好地解决这一问题。

(3) 基因诊断　基因芯片可以对遗传信息进行快速准确的分析，因此它在疾病的分子诊断中的优势是不言而喻的。从正常人的基因组中分离出DNA与DNA芯片杂交就可以得出标准图谱。从病人

的基因组中分离出DNA与DNA芯片杂交就可以得出病变图谱。通过比较、分析这两种图谱，就可以得出病变的DNA信息。

(4) 药物筛选　基因芯片技术可用于大规模药物筛选，其通用性强，并能够从基因水平解释药物的作用机理，即可以利用基因芯片分析用药前后机体的不同组织、器官基因表达的差异。

(5) 大规模DNA测序　基因芯片利用固定探针与样品进行分子杂交产生的杂交图谱而排列出待测样品的序列，这种测定方法快速，具有十分诱人的前景。

二、蛋白质芯片

蛋白质芯片又称蛋白质微阵列，是指固定于支持介质上的蛋白质构成的微阵列，它是在生物功能基因组学研究中作为基因芯片功能的补充发展起来的。与基因芯片相似，蛋白质芯片也是在一个基因芯片大小的载体上，按使用目的的不同，点布相同或不同种类的蛋白质，然后再让其与荧光标记的蛋白质特异性结合，通过扫描仪读出荧光强弱，最后由计算机分析出样本结果。理论上，蛋白质芯片可以对各种蛋白质、抗体以及配体进行检测，可弥补基因芯片检测的不足，它不仅适合于抗原、抗体的筛选，也可用于受体配体相互作用的研究。

第十节　基因组学研究概述

基因组学(genomics)这一概念最早由美国科学家Thomas Roderick于1986年提出，是指对一个基因组内的所有基因进行作图，并进行碱基序列分析以及基因定位和基因功能分析的一门学科。它包括2个方面的内容：以全基因组测序为目标的结构基因组学(structural genomics)和以基因功能鉴定为目标的功能基因组学(functional genomics)，后者又称为后基因组学。

一、结构基因组学

结构基因组学代表基因组分析的早期阶段，以建立生物体高分辨率遗传、物理和转录图谱为主，通过基因作图、碱基序列分析来确定基因组成和基因定位。由于染色体DNA不能直接进行测序，需要将基因组分解成为较易操作的小的结构区段，这个过程就是基因作图。根据使用的标记和手段的差别，基因作图可得到3类图谱，即遗传连锁图谱、物理图谱和转录图谱。

(一) 遗传连锁图谱

通过遗传重组所得到的基因线性排列图为遗传连锁图。它是通过计算连锁的遗传标记之间的重组频率，来确定它们之间的相对距离。绘制遗传连锁图的方法有很多，早期使用的多态性标记有：限制性酶切片段长度多态性(RFLP)、随机引物扩增基因组DNA(RAPD)、扩增片段长度多态性(AFLP)、20世纪80年代后出现的有短串联重复序列(STR)DNA遗传多态性分析和90年代发展的单核苷酸多态性分析(SNP)。

(二) 物理图谱

物理图谱是利用RE将染色体DNA切成片段，再根据重叠序列把片段拼接成完整的染色体，还需要确定遗传标记之间的物理距离和各种DNA序列在染色体上的准确定位。以人类基因组物理图谱为例，它包括2层含义，一是获得分布于整个基因组30 000个序列标志位点(sequence tag site，STS)；二是在此基础上构建覆盖每条染色体的大片段。

在对一段DNA序列进行染色体定位的时候，经常使用一种叫荧光原位杂交(fluorescence in situ hybridization，FISH)的技术。其基本原理是根据荧光标记的核酸探针能否与染色体DNA上的序列发生杂交，来检测特定的DNA序列是否存在以及存在的位置。其基本操作流程是(图44-22)：首先将染

色体固定到载玻片上,然后使用荧光标记的核酸或半抗原或生物素标记的核酸作为探针,在进行变性处理以后,直接与已变性的染色体上的同源互补的序列进行杂交。如果探针已被荧光标记,就可以在杂交以后就通过荧光监测系统进行观察。如果探针是用半抗原或生物素标记的,那就要使用荧光标记的抗体或亲和素,对探针上的半抗原或生物素进行追踪和定位,然后再通过荧光监测系统进行观察。

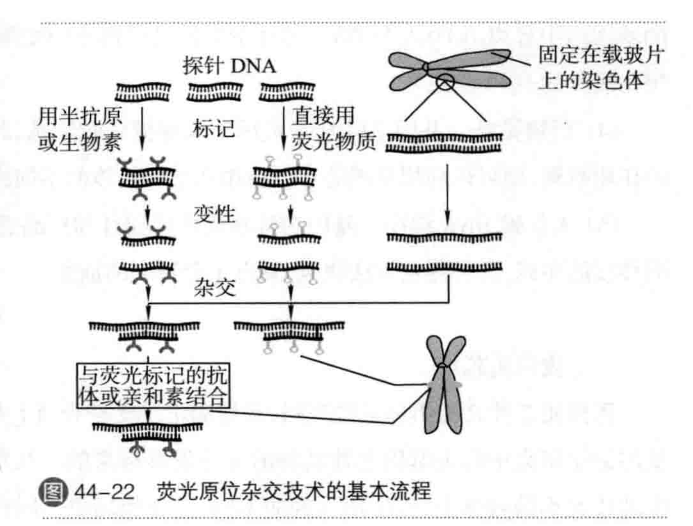


图 44-22 荧光原位杂交技术的基本流程

(三) 转录图谱

利用表达序列标签(EST)作为标记所构建的分子遗传图谱称为转录图谱。EST 是指从 cDNA 文库中随机挑取的一个克隆进行单一测序所获得的 cDNA 在 5′或 3′端的部分序列,其代表的是一个完整基因的一小部分,长度在 300~500 bp。一般而言,mRNA 的 3′UTR 代表的是每个基因比较特异的序列,将对应于 3′UTR 的 EST 序列进行放射杂交(radiation hybrid,RH)定位,即可构成由基因组成的 ETS 图。这些 EST 不但为基因组遗传图谱的构建提供了大量的分子标记,而且来自不同组织和器官的 EST 也为基因的功能研究提供了有价值的信息。此外,EST 还为基因的鉴定提供了候选基因。其不足之处在于,通过随机测序有时难以获得那些低丰度表达的基因,以及那些在特殊环境条件下诱导表达的基因。

二、功能基因组学

随着 1990 年人类基因组计划(HGP)的实施并取得巨大成功,以及其他几种模式生物(model organism)基因组计划的完成,基因组学的研究重心开始从揭示生命的所有遗传信息转移到在分子整体水平对功能的研究上。转移的主要标志是功能基因组学的产生与蛋白质组学的兴起。

功能基因组学是利用结构基因组学提供的信息来研究基因功能,它以高通量、大规模实验方法以及计算机统计分析为特征,在基因组或系统水平上全面分析基因的功能,使得生物学研究从对单一基因或蛋白质的研究转向对多个基因或蛋白质同时进行系统的研究。研究内容包括基因功能发现、基因表达分析和突变检测。采用的手段包括经典的减法杂交、差示筛选、cDNA 差异分析以及 mRNA 差异显示等,以及一些新的技术手段,例如,基因表达的序列分析(serial analysis of gene expression,SAGE)、基因芯片和蛋白质芯片等。

鉴定一个基因功能最有效的方法是观察该基因的表达被阻断或增强后在细胞和整体水平所发生的表型变化,主要手段有基因敲除、基因敲减和显性负性突变(dominant negative mutation)等技术。还需要建立模式生物体,并进行比较基因组学(comparative genomics)的研究。其中显性负性突变技术是通过基因转移,将突变的蛋白质大量引入到在特定细胞内表达,以阻断正常蛋白质的功能,从而推断野生蛋白质的功能。利用模式生物基因组与人类基因组之间编码顺序上和结构上的同源性,有助于克隆人类疾病基因,揭示基因功能和疾病分子机制,阐明物种进化关系及基因组的内在结构。比较基因组学是在基因组图谱和测序基础上,对已知的基因和基因组结构进行比较,来了解基因的功能、表达机理和物种的进化规律。

三、与基因组学相关的学科

随着基因组学研究的不断深入以及该学科对其他学科研究的渗透,一些新型的交叉学科应运而生,如营养基因组学(nutritional genomics)、环境基因组学(environmental genomics)、药物基因组学

(phamarcogenomics)、生殖基因组学(reproductive genomics)、群体基因组学(population genomics)、代谢组学(metabonomics)等和生物信息学(bioinformatics)。

框 44–2 生化热点——千人基因组计划

千人基因组计划(1 000 genomes project)是继人类基因组计划(HGP)之后又一以我们人类自身为对象的国际性重大合作研究计划,可以说是 HGP 的延续和发展,开始于 2008 年 1 月。它主要依托中国深圳华大基因研究院、英国 Sanger 研究所和美国国立卫生研究院(NIH)下属的国立人类基因组研究所(National Human Genome Research Institute,NHGRI)来实施。通过该计划,科学家们打算在三年内,借助目前几种新的高通量测序平台,测定出至少一千名匿名者的全基因组序列,并在此基础上建立最详细的人类遗传变异(human genetic variation)目录,即绘制迄今为止最详尽、最有医学应用价值的人类基因组遗传多态性图谱。这些匿名者具有不同的种族背景,测序的人群包括:尼日利亚伊巴丹区域的约鲁巴人;东京的日本人;北京的中国人;美国犹他州的北欧和西欧人后裔、休斯敦的 Gujarati 印第安人、丹佛的中国人、洛杉矶的墨西哥人后裔以及西南部的非洲人后裔;肯尼亚 Webuye 的 Luhya 人和 Kinyawa 的 Maasai 人;意大利的 Toscani 居民。

深圳华大基因研究院作为发起单位之一,不仅承担了 400 个黄种人全基因组样本的测序和分析工作,还帮助完成了非洲人群的全部测序和分析任务。迄今为止所获得的全部数据已存储于该计划所设立的公共数据库(http://www.1000genomes.org/data 中有链接),以供公众免费获取和浏览。该数据库最终将包含来自全球 27 个族群的 2 500 个人的全部基因组信息。这一数据库的建立可为各种疾病的关联分析提供详细的基础数据,并为解释人类重大疾病发病机理以及开展疾病的个性化预测、预防和治疗奠定坚实的基础。此外,它还可加深人们对人类群体遗传学的理解,促进人类进化史的研究。

生物信息学是以生物大分子为研究对象,以计算机为辅助工具,运用数学和信息学的观点、理论和方法去研究生命现象,组织和分析呈指数级增长的生物信息数据的一门科学。研究重点集中在基因组学和蛋白质组学两个方面。其关注的研究热点包括:序列比对;基因识别和 DNA 序列分析;蛋白质结构预测;分子进化等。

科学故事——PCR 的发明

Louis Pasteur 曾经说过:"机会只青睐那些有准备头脑的人"。科学发展史上有许多例子可以验证这句名言,例如牛顿从树上掉落下来的苹果发现了重力,弗莱明在污染的培养皿上发现了青霉素,而 PCR 的发现又是一个十分有趣的例子。

PCR 的发明人,一般公认是美国科学家 Kary Mullis,他也因此获得了 1993 年的诺贝尔化学奖。Mullis 曾在 1990 年 *Scientific American* 杂志上发表的一篇文章以及 1998 年的自传《心灵裸舞》(Dancing Naked in the Mind Field)中,提到 PCR 这个构想的灵感。

1979 年,Mullis 离开先前工作的美国加利福尼亚大学旧金山分校的医学院,进了一家叫 Cetus 的私人生物技术公司任职。当年,生物技术公司还处于萌芽阶段,很少有学术界人士愿意离开大学到私人企业工作。就算是到有规模的大药厂,同样也得不到多数同行的认同,反而被认为是学术生涯的终点。然而,Cetus 公司却不一样,这家公司在自由开放的风气下,集结了一批有能力、有梦想的科学家,朝着既定的目标共同前进。这和一般大学里各大教授及实验室的主持人关起门来各行其是的做法相当不同。Cetus 聘用 Mullis,是看重他有机化学合成的专长,负责合成寡聚核苷酸,以供实验所用。

其实,Mullis 在 Cetus 的工作,算不上什么研究,只是设法改进寡聚核苷酸合成的效率而已。1981 年,他被提升为寡核苷酸合成部门的主管。为了提高产量和节省时间,他省去了质量控制这一环节,引起使用单位的不满,抱怨质量不佳的寡聚核苷酸使得他们的研究出现问题,Mullis 则回击道,是使用单位本身的

能力不足所致。

PCR 的构想，也就是在这样的背景下诞生的。根据 Mullis 自己的说法，他本来并不是想扩增微量的 DNA，而只想使用染色的方法鉴定一个突变的基因。这是一个失败的想法，但却把他引向 DNA 的扩增。一直在做突变检测工作导致 Mullis 想到细胞内 DNA 复制的反应。碰巧，Mullis 那时一直沉迷于计算机如何处理指数功能，这导致他将两个事情联系在一起。每一个分子生物学家都知道 DNA 复制反应(加倍)，而且知道指数函数。但只有 Mullis 同时想到两件事情，从而产生了指数扩增 DNA 的构想。Mullis 意识到，少量的东西如果持续加倍，就会增加得很快。例如，将一个特定的 DNA 片段加倍 30 次，则会产生 2^{30} 拷贝，这么多拷贝的 DNA 足以让研究者进行相关的科学研究。这种想法有一天终于得以爆发！那是在 1983 年春天的一个星期五的晚上，Mullis 开着银色的本田 Civic，带着女友前往乡间的小屋度周末。在北加州蜿蜒的 128 高速公路上开着车，一段 DNA 反复复制的景象，在他的脑海里冒了出来。于是他停下了车，叫醒了正在熟睡的女友，激动地解释起他的想法。Mullis 原以为这样简单的想法，应该有人提出过，但搜索文献后却发现没有。在"猛然顿悟"之后的 3~5 个月间，Mullis 并没有任何行动，原因如今也不清楚。同年 8 月，Mullis 首次在公司里正式作了有关 PCR 原理的报告，听者反应冷淡。一来，大家已经习惯了他的胡思乱想；再者，多数人认为，这个原理太简单了。如果可行的话，就一定早有人做过。否则，一定有什么不可行之处，但也没有人明确说得出来，为什么不可行。

于是，Mullis 开始着手证明这个构想的可行性。从 1983 年 9 月起，Mullis 陆续进行了一些实验，换过几种 DNA 模板，也尝试过不同的升温、降温周期，结果都不理想，顶多只在电泳凝胶上发现一个若隐若现的条带，无法说服别人 PCR 起了扩增的功效。1984 年 6 月，Mullis 在公司又因男女关系惹出事端，引起众怒，濒临被开除的命运。结果是引荐他进入公司的上司为他求情，只免除了他的主管职务，同时限定他在一年内把 PCR 方法建立起来。

任何研究方法从概念的提出发展到实际应用阶段，所需投入的精力与时间大多被一般人所低估。由于 Mullis 以前没有接受过分子生物学的训练，公司派了技术员协助，前后一共有三位。这些人在 PCR 的发展上，发挥了重要的作用。1984 年 11 月，Mullis 的技术员首次取得可信的结果，证明了 PCR 的可行性。于是在 1985 年初，公司决定又派上了技术精湛的日裔技术员 Randall Saiki，这是一项正确的决定。在自动化的仪器出现之前，PCR 是一项十分繁琐的技术，同时需要许多人在一堆试管、秒表和不同温度的水浴锅中忙个不停，甚至还要独立的作业空间防止可能的污染，这需要长时间的反复操作，手脚不利落的人是做不来的。Saiki 的结果则干净漂亮，让人心服口服。

到了 1985 年春天，Cetus 的高级主管已经对 PCR 的潜力深信不疑，也开始担心消息外泄，而让其他公司取得先机。同年 3 月，他们递出了第一个专利申请，也准备在 10 月份举行的美国遗传学会年会上报告成果，但之前必须将正式的论文写好寄出后才保险。他们决定撰写两篇论文，一篇关于 PCR 的原理，由 Mullis 执笔先行发表，第二篇则集中在 PCR 的应用上，以 Saiki 的实验结果为主，随后推出。结果整个夏天，Mullis 都在玩电脑，一再拖延论文的写作。到 9 月下旬有关 PCR 应用的文章写好寄出时，Mullis 还没有动静。因此，第一篇提到 PCR 这个方法的论文，于 1985 年 12 月 20 日发表在 *Science* 上，共有七位作者，Saiki 为第一作者，Mullis 则排第四。

直到 12 月，Mullis 才将论文写好，并投给 *Nature*。但 Mullis 忘记了附上一封给编辑的信，当然也就没有说明该文与 *Science* 上的那篇论文有何不同，结果被拒。震惊之余，他转投 *Science*，并由 Cetus 的主管帮助写了封信给编辑，结果仍然被拒。这时，Mullis 把怨气转向公司，认为那是公司的阴谋，想要窃取他发明 PCR 的功劳。实际上，PCR 的概念源自 Mullis，对此没有什么人持异议，只不过将概念付诸实施的过程，就复杂得多了。

Mullis 的文章两度被拒以后，公司里有人建议投给 *Methods of Enzymology*，主要是因为有人与该刊主编吴瑞相熟，容易沟通，同时 PCR 的性质也适合强调方法学的刊物。于是，Mullis 的文章终于得到发表，只不过整整晚了一年，到 1987 年初才问世。

为了表示他们并无意争功，Cetus 的主管向冷泉港实验室的 Watson 推荐，让 Mullis 在于 1986 年 5 月举行的“人类分子生物学”专题研讨会中，报告 PCR 的原理及实际应用结果。这是 Mullis 生平第一次受邀演讲，分子生物学界许多著名学者也都在场。结果 Mullis 表现不错，建立了往后人们的印象：PCR 是他一手发明的。冷泉港专题研讨会的专刊于 1986 年底出版，Mullis 挂了头名。

自此，PCR 之名及其强大的应用性就广为人知了。然而，将 PCR 变成真正成熟技术的“临门一脚”，则是耐高温 DNA 聚合酶的引进。

在 PCR 的操作过程中，需要进行反复升温与降温的循环步骤，而前一次循环所使用的大肠杆菌 DNA 聚合酶在高温下就变性失活了，因此在每一次循环反应之后，都要加入新的聚合酶。这种做法既烦琐，又昂贵。按当时的价格，一次循环所需的聚合酶值 1 美元，30 个循环下来就是 30 美元，循环更多次就不用说了。因此，1986 年春，Mullis 首次提出使用耐高温酶的想法。经过文献搜索，果然找到了 2 篇有关文献，较早的一篇是在美国做的，另一篇则是俄罗斯科学家的成果。

第一篇报道分离耐高温 DNA 聚合酶的工作，是一位来自台湾的年轻科学家钱嘉韵。1973 年，钱嘉韵到美国俄亥俄州的辛辛那提大学生物系就读。她的导师 J. Trela 对一种在黄石公园温泉里发现的水生嗜热菌（*Thermus aquaticus*，*Taq*）感到好奇，就让钱及另一位美国学生以该细菌作为论文研究的题目。在另一位老师的指导下，钱嘉韵成功地从该细菌中，分离出耐高温的 *Taq* DNA 聚合酶，并将研究成果发表在 1976 年的 *Journal of Bacteriology* 杂志上。

Mullis 虽然提出将 *Taq* DNA 聚合酶应用到 PCR 的建议，但当时并没有现成的酶可用，他得想办法自己分离。Cetus 有全套分离蛋白质的设备，也有人愿意指导，但 Mullis 却并不急。等了几个月后，公司其他人只有自己动手，按着先前钱嘉韵等人发表的步骤，三个星期就纯化出 *Taq* DNA 聚合酶。1986 年 6 月，Saiki 首度将其应用于 PCR，效果就好得惊人，可以说是一炮打响。*Taq* DNA 聚合酶不但大大简化了 PCR 工作，而且其专一性及活性都比之前使用的大肠杆菌 DNA 酶更强，背景杂带也几乎都消除了。自此，PCR 大获成功！

e44-8 本章小结填填看

推荐网址：

1. http://en.wikipedia.org/wiki/Molecular_cloning（维基百科有关分子克隆的内容）
2. http://www.cellbio.com/protocols.html（在线的细胞和分子生物学实验技术）
3. http://www.protocol-online.org/prot/Molecular_Biology/（在线的分子生物学实验技术）

参考文献：

1. Pfeifer A, Verma I M. Gene therapy: promises and problems. *Annu Rev Genomics Hum Genet*, 2: 177–211, 2001.
2. Yeates T O, et al. Protein function in the post-genomic era. *Nature*, 405, 823–826, 2000.
3. Pandey A, Mann M. Proteomics to study genes and genomes. *Nature*, 405: 837–846, 2000.
4. Brown P O, Botstein D. Exploring the new world of the genome with DNA microarrays. *Nat Genet*, 21: 33–37, 1999.
5. Arnheim N, Erlich H. Polymerase chain reaction strategy. *Annu Rev Biochem*, 61: 131–156, 1992.
6. Lobban P E, Kaiser A D. Enzymatic end-to-end joining of DNA molecules. *J Mol Biol*, 78: 453–471, 1973.
7. Berg P, et al. Biochemical method for inserting new genetic information into DNA of simian virus 40: circular SV40 DNA molecules containing lambda phage genes and the galactose operon of *Escherichia coli*. *Proc Natl Acad Sci*, 69: 2 904–2 909, 1972.

分子生物学思考题

1. 大肠杆菌的 DNA 聚合酶 I 具有 3 种不同的酶活性，这些酶活性对 DNA 复制的忠实性是不是都有贡献？为什么？

2. 以下几种情况下，细胞要么立刻停止复制 DNA（快停），要么只能停止下一轮 DNA 复制（慢停）。试预测每一种情况下究竟是快停还是慢停？为什么？假定细胞并没有进行同步化处理，即每一个细胞可能处在细胞周期的不同阶段。

(1) 一种大肠杆菌温度敏感型突变株：其 DnaA 蛋白在 30℃下功能正常，42℃下变性。将这种大肠杆菌的培养温度从 30℃提高到 42℃。

(2) 大肠杆菌的 DNA 聚合酶Ⅲ的 γ 亚基受诱导大量表达。

(3) 一种大肠杆菌温度敏感型突变株：其 DnaB 蛋白在 30℃下功能正常，42℃下变性。将这种大肠杆菌的培养温度从 30℃ 提高到 42℃。

(4) 含有正常 CDC6 的酵母细胞被诱导大量表达不能被 CDK 磷酸化的 CDC6 变体。

3. 2- 氨基嘌呤（2-AP）是 A 的类似物。在细菌体内它可转变成相应的脱氧核苷三磷酸，并在 DNA 复制中参入到子链之中，引发突变。2-AP 一般导致的突变为碱基转换，为什么？尽管 2-AP 被认为只能导致碱基转换，但将它加入到细菌培养基上以后，发现细菌体内发生的移码突变也有显著的提高，这又是为什么？你如何设计一个实验以支持你的解释？

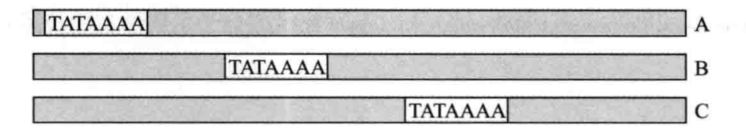
4. RecA 蛋白无活性的大肠杆菌突变株（$recA^-$）在受到许多诱变因素（如离子辐射或 UV 辐射）作用后，发生突变的效率明显低于野生型大肠杆菌（$recA^+$），为什么？

5. 有人构建了一个受到 UV 辐射损伤的 *lacZ* 基因，并将其导入到大肠杆菌细胞内。下图链上的 X 代表嘧啶二聚体。A 和 B 的差别是在于：底物 A 的损伤位于 *lacZ* 基因的编码链上，B 的损伤位于 *lacZ* 基因的模板链上。试问在乳糖操纵子受到诱导之后，哪一个底物上的嘧啶二聚体更容易被修复？为什么？

X
5′
3′
A

5′
3′
X
B

6. 如果将以下 3 种不同的 DNA 片段与 TBP 结合，然后在凝胶上观察它们的电泳行为，你认为哪一个片段在结合 TBP 后跑得最快？哪一个跑得最慢？为什么？



7. 在体外将 RNA 聚合酶全酶与松弛型质粒保温在一起，可以导致 DNA 形成超螺旋吗？如果能的话，是正超螺旋还是负超螺旋？

8. 如果使用基因工程的手段，将 RNA 聚合酶的 σ 因子与核心酶的某一个亚基融合在一起，形成融合多肽（假定融合对各个亚基原来的结构没有影响），预测这样的全酶对基因的转录会有什么影响？为什么？

9. His 784 位于 RNA 聚合酶的活性中心，它在所有的 RNA 聚合酶中都是保守的，但很少存在于各种 DNA 聚合酶。基于此，你认为 His 784 在 RNA 聚合酶中行使何种功能？

10. 有人发现，一种特殊的真核 RNA 病毒能从一个基因编码出一长一短的两种 mRNA。分析它们的翻译产物，发现它们翻译出来的两条多肽链在 N 端具有相同的氨基酸序列，但 C 端不一样，更另人奇怪的是，两条多肽中较长的一条由短的 mRNA 所编码。试对这些现象给予合理的解释。

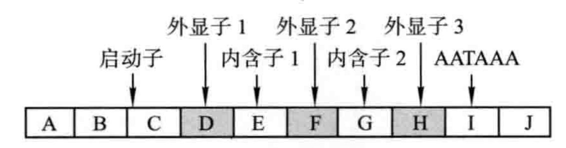
11. 下图显示的是果蝇的一种蛋白质基因的结构，它共由 10 个片段组成，分别以 A~J 表示。

(1) 哪些片段会出现在初级转录物？

(2) 哪些片段出现在成熟的 mRNA 分子上？

(3) 成熟的 mRNA 含有哪些不是由基因编码的信息？

(4) 如果以成熟的 mRNA 做探针进行 R- 环实验，试画出最后实验结果的草图。



12. RNA 聚合酶Ⅱ催化亚基的 CTD，含有多个重复的模体序列 YSPTSPS。削短一些重复序列可导致细胞出现某些严重的缺陷。预测细胞内的以下过程受到何种影响？

①rDNA 的转录。②mRNA 的加尾。③介导因子对转录的激活。④SnRNA 的转录。⑤mRNA 的稳定性。

13. 在过量表达 λ 噬菌体 *int* 基因的时候，Ken Zahn 发现一个很有趣的现象：他将 *int* 基因放在一个表达质粒上，使其位于一个可调控的强启动子的下游，且将编码 int 的 mRNA 带上强的核糖体结合位点。当 *int* 受诱导表达的时候，宿主细胞的蛋白质合成受到抑制，细胞最终死亡。Zahn 很快又发现，细胞死亡并不是 Int 蛋白的毒性引起，因为 *int* 的过量表达产生的效应可以通过导入一个表达 tRNAArg 的基因而逆转。这种 tRNAArg 阅读 AGA 密码子。

(1) 你如何解释以上这种有趣的现象？

(2) 你能否想出一种办法让 *int* 过量表达，但细胞不会死亡？

14. 在大肠杆菌细胞里表达来自嗜热菌的基因现在已经十分普遍。然而，有时表达的效果不错，有时表达的则不理想。Ishida 和 Oshima 在研究中发现，那些表达量极低的基因在紧靠 RBS 附近含有回文序列。如果将此序列去除，这些基因在大肠杆菌中的表达量将大大提高。为什么这种回文序列在嗜热菌体内不影响基因的表达，但在大肠杆菌中却能降低基因表达的效率？

15. 在细菌翻译的时候，氨基酸错误参入的概率是 10^{-3}。如果一种对细胞极为重要的蛋白质在翻译的时候，在一个关键的位置参入了一个错误的氨基酸，你认为这会对细菌带来何种后果？为什么？

16. 通常细胞质里只有一种 tRNA 运载 Tyr。你认为这种 tRNA 的反密码子是什么？

17. 细菌的核苷酸还原酶由两个不同的亚基组成(RR-A 和 RR-B)。在生长的细菌细胞内，RR-A 和 RR-B 通常一直表达，两者的表达水平相同。然而，如果 DNA 复制受阻，dTTP 浓度升高，那么这两个亚基的表达开始下降。

(1) 提出一种可能的机制，解释细菌内 RR-A 和 RR-B 的表达是如何协调一致的。

(2) 细菌的核苷酸还原酶基因表达的最高水平本来就低，即使在 dTTP 浓度不高的时候。根据细胞内 RNA 表达的量，哪一个基因表达的水平最高？其中的原因是什么？

(3) 你如何解释细菌核苷酸还原酶基因的表达先天不高这一现象(不考虑 dTTP 浓度对基因表达的影响)。

18. 有人在研究组蛋白密码时，将酵母 H3 组蛋白的尾巴缺失，发现这对酵母的生长没有什么影响。然而，如果将酵母 H3 和 H4 的尾巴都缺失，酵母即死亡。对此现象你如何解释？

19. 许多基因表达调控蛋白在结合 DNA 以后，导致 DNA 在结合位点附近产生弯曲。测定 DNA 弯曲的方法是一种特殊的凝胶电泳分析。这种分析方法是建立在以下现象的基础之上：一个中间发生弯曲的 DNA 的泳动速率，要低于没有弯曲或在一侧出现弯曲的 DNA。如果给你 T4 DNA 连接酶，你能否想出一种更好的方法来测定 DNA 的弯曲？

20. 有人使用 PCR 扩增一个 2 kb 长的基因，他在准备扩增反应的时候，犯了一些错误。根据你学过的分子生物学知识，你认为这些错误对 PCR 的结果有何影响？他是否需要重新准备反应系统？

(1) 在反应系统中，他除了添加 dNTP，还加入了少量的 ddNTP。

(2) DNA 模板的量比原计划多加了两倍。

(3) 他使用了大肠杆菌的 DNA 聚合酶代替 *Taq*DNA 聚合酶。

(4) 他设置的扩增循环条件是：1 min @ 37℃，1 min@ 50℃，1 min @72℃。

(5) 他使用的引物序列与基因的编码链两侧的序列完全相同。

(6) 他使用 Mn^{2+} 代替了 Mg^{2+}。

主要参考书目

1. 王镜岩,朱圣庚,徐长法 . 3 版 . 北京:高等教育出版社,2002.
2. 杨荣武等 . 分子生物学 . 南京:南京大学出版社,2007.
3. 许根俊 . 酶的作用机理 . 北京:科学出版社,1983.
4. 朱玉贤 . 现代分子生物学 . 3 版 . 北京:高等教育出版社,2007.
5. 王金发 . 细胞生物学 . 北京:科学出版社,2003.
6. 王德宝,祁国荣 . 核酸结构、功能与合成 . 北京:科学出版社,1987.
7. 历朝龙 . 生物化学与分子生物学 . 北京:中国医药科技出版社,2001.
8. 罗纪盛等 . 生物化学简明教程 . 3 版 . 北京:高等教育出版社,2006.
9. Nelson D L, Cox M M. Lehninger principles of biochemistry. 5th ed. W.H. Freeman and Company, 2008.
10. Moran L A, Horton R A, Gray Scrimgeour and Marc Perry. Principles of Biochemistry. 5th ed. Pearson Education Inc, 2011.
11. Berg J M, Tymoczko J L, Stryer L. Biochemistry. 6th ed. W.H. Freeman and Company, 2006.
12. Campbell M K, Farrell S O. Biochemistry. 7th ed. Brooks Cole, 2011.
13. Garrett R H, Grisham C M. Biochemistry. 4th ed. Brooks Cole, 2010.
14. Denniston K J, Topping J J, Caret R L. General, Organic and Biochemistry. 6th ed. McGraw-Hill Companies, 2008.
15. Watson J D, Baker T A, Bell S P, et al. Molecular Biology of the Gene. 6th ed. Benjamin Cummings, 2007.
16. Krebs J E, Goldstein E S, Kilpatrick S T. Lewin's Genes X. Pearson Education Inc, 2009.
17. Weaver R F. Molecular Biology. 5th ed. McGraw-Hill Companies, 2011.
18. Alberts B, Johnson A, Lewis J, et al, Molecular Biology of the Cell. 5th ed. Garland Science, 2007.
19. Russell P J. iGenetics: a molecular approach. 3rd ed. Pearson Education Inc, 2010.
20. Voet D, Voet J G, Pratt C W. Fundamentals of Biochemistry. 2nd ed. John Wiley & Sons Inc, 2006.
21. Whitford D. Proteins: structure and function. John Wiley & Sons Inc., 2005.
22. Comperts B D, Kramer I M, Peter E R. Tatham. Signal Transduction. 2nd ed. Elsevier Inc, 2009.
23. Pierce B A. Genetics Essential. W.H. Freeman and Company, 2010.
24. Dale J W, Park S F. Molecular Genetics of Bacteria. 5th ed. John Wiley & Sons Ltd, 2010.
25. Sambrook J, Russell D W. Molecular Cloning: A laboratory Manual. 3rd ed. Cold Spring Harbor Laboratory Press, 2001.

推荐网站

1. http://js.nclass.org/vc/802276/(南京大学生物化学国家级双语教学示范课程网站,内有各种辅助教学和学习的资料)
2. http://www.wiley.com/legacy/college/boyer/0470003790/animations/animations.htm(一个学习生物化学的动画互动网站)
3. http://www.freebooks4doctors.com/(免费的生命科学和医学电子书籍链接)
4. http://www.avaxhome.ws/ebooks/(免费的生命科学和医学电子书籍链接)
5. http://www.bioon.com/(生物谷——一个生物医学门户网站)
6. http://www.dxy.cn/bbs/(丁香园——医学、药学、生命科学专业知识检索与交流论坛)
7. http://www.researchchannel.org/prog/(免费的生命科学和医学在内多种学科的讲座视频)
8. http://www.sciencedaily.com/(每日提供最新的科技新闻,许多与生命科学相关)
9. http://bcs.whfreeman.com/lehninger/default.asp(Lehninger 编著的《生物化学原理》配套的网站)
10. http://www.pearsonhighered.com/mathews/(Mathews 编著的《生物化学》配套的网站)
11. http://www.web-books.com/MoBio/(免费的分子生物学教材)
12. http://www.ncbi.nlm.nih.gov/books/NBK21154/(Stryer 编著的《生物化学》第五版)
13. http://themedicalbiochemistrypage.org/(免费的医学生物学教材)
14. http://www.51qe.cn/book/book18.php/(免费的分子生物学中文教材)
15. http://www.ncbi.nih.gov/(美国生物技术信息中心(NCBI))
16. http://www.protocol-online.org/(提供各种生命科学实验方法)
17. http://cmbi.bjmu.edu.cn/(中国医学生物信息网)
18. http://bingle.pku.edu.cn/(北京大学天网 FTP、WWW 文件搜索引擎)
19. http://www.science.com/(科学杂志)
20. http://www.nature.com/(自然杂志)
21. http://www.cell.com/(细胞杂志)
22. http://www.pnas.com/(美国科学院院刊)
23. http://en.wikipedia.org/(维基百科)
24. http://au.expasy.org/(蛋白质序列数据库)
25. http://genomesonline.org/(各种生物的基因组数据)

本书常见英文缩写

缩写	英文全称	中文译称
aaRS	aminoacyl-tRNA synthetase	氨酰 tRNA 合成酶
Ab	antibody	抗体
AC	adenylate cyclase	腺苷酸环化酶
ACP	acyl carrier protein	脂酰基载体蛋白
ACAT	acyl-CoA:cholesterol acyl transferase	脂酰 -CoA:胆固醇脂酰基转移酶
ACTH	adrenocorticotropin	促肾上腺皮质激素
AD	Alzheimer's disease	阿尔茨海默氏病
ADA	adenosine deaminase	腺苷脱氨酶
ADH	antidiuretic hormone	抗利尿激素
Ag	antigen	抗原
AIDS	acquired immune deficiency syndrome	获得性免疫缺陷综合征
AMPK	AMP-activated protein kinase	AMP 激活的蛋白质激酶
ANF	atrial natriuretic factor	心钠素或心房利钠因子
APRT	adenine phosphoribosyltransferase	腺嘌呤磷酸核糖转移酶
AQP	aquaporin	水通道蛋白
ARS	autonomously replicating sequence	自主复制序列
ATC	aspartate transcarbamoylase	天冬氨酸转氨甲酰酶
AZT	azidothymine	叠氮胸苷(齐多夫定)
BCCP	biotin carboxyl carrier protein	生物素羧基载体蛋白
BER	base excision repair	碱基切除修复
BPG	2,3-bisphosphoglycerate	2,3- 二磷酸甘油酸
cAMP	cyclic AMP	环腺苷酸
CaMV	Cauliflower mosaic virus	花椰菜镶嵌病毒
cccDNA	covalently-closed circular DNA	共价闭环 DNA
CCK	cholecystokinin	胆囊收缩素
CDK	cyclin-dependent kinase	依赖于周期蛋白的蛋白质激酶
cDNA	complementary DNA	互补 DNA
cGMP	cyclic GMP	环鸟苷酸
CGRP	calcitonin gene-related peptide	降钙素基因相关肽
CJD	Creutzfeldt-Jakob disease	克雅氏病
CLIP	corticotrophin-like intermediate peptide	促皮质激素类中间肽
CM	chylomicron	乳糜微粒
CoA	coenzyme A	辅酶 A
CoQ	coenzyme Q	辅酶 Q
CPA	carboxypeptidsase A	羧肽酶 A

CPS	carbamyl phosphate synthetase	氨甲酰磷酸合成酶
CPT	carnitine palmitoyltransferase	肉碱－软脂酰转移酶
CRE	cAMP-response element	cAMP 应答元件
CRP	cAMP receptor protein	cAMP 受体蛋白
CSF	colony stimulating factor	集落刺激因子
CTD	carboxyl-terminal domain	羧基端结构域
DAG	diacyl glycerol	甘油二酯
DCCD	dicyclohexylcarbodiimide	二环己基碳二亚胺
ddNTP	3′,5′-dideoxynucleotide	3′,5′－双脱氧核苷酸
DHA	docosahexaenoic acid	二十二碳四烯酸(脑黄金)
DHAP	dihydroxyacetone phosphate	磷酸二羟丙酮
DHEA	dehydroepiandrosterone	脱氢异雄酮
DHT	dihydrotestosterone	二氢睾酮
DRM	detergent-resistant membrane	抗去垢剂膜
DIPF	diisopropylfluorophosphate	二异丙基氟磷酸
DMS	dimethylsulphate	硫酸二甲酯
DMT	dimethoxytrityl group	二甲氧基三苯基
DNA	deoxynucleic acid	脱氧核糖核酸
DNFB	2,4-dinitrofluorobenzene	2,4－二硝基氟苯
DNP	2,4-dinitrophenol	2,4－二硝基苯酚
DTT	dithiothreitol	二硫苏糖醇
EB	ethidium bromide	溴乙锭
ECM	extracellular matrix	细胞外基质
EGF	epidermal growth factor	表皮生长因子
EFA	essential fatty acid	必需脂肪酸
EMP	Embden-Meyerhof pathway	Embden-Meyerhof 途径
EMSA	electrophoretic mobility shift assay	电泳泳动变化分析
EPO	erythropoietin	红细胞生成素
ERAD	ER-associated degradation pathway	ER 相关的蛋白质降解途径
ESE	exonic splicing enhancer	外显子剪接增强子
ETS	electron transport system	电子传递体系
EST	expression sequence tag	表达序列标签
FAD	flavin adenine dinucleotide	黄素腺嘌呤二核苷酸
FGF	fibroblast growth factor	成纤维细胞生长因子
FMN	flavin mononucleotide	黄素单核苷酸
FMOC	9-fluorenylmethoxycarbonyl	9－芴甲氧羰基
FSH	follicle-stimulating hormone	卵泡刺激素
GABA	gamma-aminobutyric acid	γ－氨基丁酸
GAG	glycosaminoglycan	糖胺聚糖
Gal	galactose	半乳糖
GAP	GTPase activating protein	GTP 酶激活蛋白
GAPDH	glyceraldehyde-3-phosphate dehydrogenase	3－磷酸甘油醛脱氢酶

GC	guanylyl cyclase	鸟苷酸环化酶
GEF	guanine nucleotide exchange factor	鸟苷酸交换因子
GFP	green fluorescent protein	绿色荧光蛋白
GIP	gastric inhibitory polypeptide	抑胃肽
Glc	glucose	葡萄糖
GnRH	gonandotropin-releasing hormone	促性腺激素释放激素
GPT	glutamate：pyruvate transaminase	谷丙转氨酶
GPCR	G protein-coupled receptor	与 G 蛋白偶联的受体
GOT	glutamate：oxaloacetate transaminase	谷草转氨酶
GRE	glucocorticoid hormone response element	糖皮质激素应答元件
GSK	glycogen synthase kinase	糖原合酶激酶
gRNA	guide RNA	指导 RNA
HAT	histone acetyltransferase	组蛋白乙酰基转移酶
HDL	high density lipoprotein	高密度脂蛋白
H-DNA	hinged DNA	铰链 DNA
HGP	Human Genome Project	人类基因组计划
HGPRT	hypoxanthine-guanine phosphoribosyltransferase	次黄嘌呤－鸟嘌呤磷酸核糖转移酶
HIC	hydrophobic interaction chromatography	疏水作用层析
HIF	hypoxia inducible factor	缺氧诱导因子
HSP	heat-shock protein	热激蛋白
HMS	hexose monophosphate shunt	磷酸己糖支路
HPLC	high-performance liquid chromatography	高效液相层析
HRE	hormone response element	激素应答元件
HSL	hormone-sensitive lipase	激素敏感性脂肪酶
HTN	hemodynamic shear stress	血液动力学微剪应力
HTLV	human T-cell leukemia virus	人 T 细胞白血病病毒
IAA	indole-3-acetic acid	吲哚乙酸
IDH	isocitrate dehydrogenase	异柠檬酸脱氢酶
IDL	intermediate density lipoprotein	中间密度脂蛋白
IEF	isoelectric focusing	等电聚焦
IGF	insulin like growth factor	类胰岛素生长因子
IgG	immunoglobulin	免疫球蛋白
IP_3	inositol-1,4,5-triphosphate	1,4,5-三磷酸肌醇
IPTG	isopropyl-D-thiogalacto-pyranoside	异丙基硫代-β-D-半乳糖苷
IRE	insulin-response element	胰岛素应答元件
IRE	iron-response element	铁应答元件
IS	insertion sequence	插入序列
IUB	International Union of Biochemistry	国际生物化学联合会
IUPAC	International Union of Pure and Applied Chemistry	国际纯粹与应用化学联合会
LCAT	lecithin：cholesterol acyl transferase	卵磷脂：胆固醇脂酰基转移酶
LDH	lactate dehydrogenase	乳酸脱氢酶

LDL	low density lipoprotein	低密度脂蛋白
LHC	light-harvesting complex	聚光复合体
LINE	long interspersed nuclear element	长散布核元件
LTR	long terminal repeat	长末端重复序列
MAR	matrix-associated region	基质相关区域
MAPK	mitogen-activated protein kinase	有丝分裂原激活蛋白激酶
MCS	multiple cloning site	多克隆位点
micRNA	mRNA-interfering complementary RNA	信使干扰互补 RNA
MMP	matrix metalloproteinase	基质金属蛋白酶
MMR	mismatch repair	错配修复
MMTV	mouse mammary tumor virus	小鼠乳腺肿瘤病毒
MRE	metal response element	金属应答元件
mRNA	messenger RNA	信使 RNA
MSH	melanocyte-stimulating hormone	促黑激素
MSF	mitochondrial-import stimulating factor	线粒体输入刺激因子
MSP	manganese stabilizing protein	锰稳定蛋白
MTX	methotrexate	氨甲蝶呤
NAG	*N*-acetylglucosamine	*N*- 乙酰葡糖胺
NAM	*N*-acetylmuramic acid	*N*- 乙酰胞壁酸
NCAM	neural cell adhesion molecule	神经细胞黏附分子
NDP	nucleoside diphosphate	核苷二磷酸
NER	nucleotide excision repair	核苷酸切除修复
NGF	nerve growth factor	神经生长因子
NHEJ	non-homologous end joining	非同源末端连接
NLS	nuclear localization signal	细胞核定位信号
NMD	nonsense-mediated mRNA decay	无义介导的 mRNA 降解
NMP	nucleoside monophosphate	核苷酸单磷酸
NMR	nuclear magnetic resonance	核磁共振
NSD	nonstop mRNA decay	无终止 mRNA 降解
NTP	nucleoside triphosphate	核苷三磷酸
OEC	oxygen-evolving complex	产氧复合物
ORC	origin recognition complex of proteins	起始区识别蛋白质复合体
ORF	open reading frame	开放阅读框
OxP	oxidative phosphorylation	氧化磷酸化
PABP	poly A binding protein	poly A 结合蛋白
PAF	platelet-activating factor	血小板激活因子
PCNA	proliferating cell nuclear antigen	分裂细胞核抗原
PCR	polymerase chain reaction	聚合酶链式反应
PD	Parkinson's disease	帕金森氏病
PDB	protein data bank	蛋白质数据库
PDE	phosphordiesterase	磷酸二脂酶
PDGF	platelet derived growth factor	血小板衍生的生长因子

PDI	protein disulphide isomerase	蛋白质二硫键异构酶
PEG	polyethylene glycol	聚乙二醇
PEP	phosphoenolpyruvate	磷酸烯醇式丙酮酸
PFK	phosphofructokinase	磷酸果糖激酶
PGTP	phosphoryl group transfer potential	磷酸基团转移势能
PIC	pre-initiation complex	预转录起始复合物
PITC	phenylisothiocyanate	苯异硫氰酸酯
PKA	protein kinase A	蛋白质激酶 A
PKB	protein kinase B	蛋白质激酶 B
PKC	protein kinase C	蛋白质激酶 C
pmf	proton motive force	质子驱动力
PLA	phospholipase A	磷脂酶 A
PLC	phospholipase C	磷脂酶 C
PLP	pyridoxal phosphate	磷酸吡哆醛
PNP	polynucleotide phosphorylase	多聚核苷酸磷酸化酶
PTH	parathyroid hormone	甲状旁腺素
PPI	peptidyl-prolyl *cis-trans* isomerase	肽酰脯氨酰顺反异构酶
PPP	phosphate pentose pathway	磷酸戊糖途径
PRPP	5-phosphoribosyl-1-α-pyrophosphate	1-α-焦磷酸-5-磷酸核糖
PTGS	post transcriptional gene silencing	转录后基因沉默
PYLIS	pyrrolysine insertion sequence	吡咯赖氨酸插入序列
RAG	recombination-activating gene	重组激活基因
RE	restriction endonuclease	限制性内切酶
RIA	radioimmunoassay	放射免疫测定法
RPA	replication protein A	复制蛋白 A
RSS	recombination signal sequence	重组信号序列
RTK	receptor tyrosine kinase	受体酪氨酸蛋白激酶
RT	reverse transcriptase	逆转录酶
RNA	ribonucleic acid	核糖核酸
RAP-PCR	RNA arbitrarily primed PCR	RNA 随机引物聚合酶链式反应
RNAi	RNA interference	RNA 干扰
rRNA	ribosomal RNA	核糖体 RNA
RSV	Rous sarcoma virus	劳氏肉瘤病毒
Rubisco	ribulose bisphosphate carboxylase/oxygenase	1,5-二磷酸核酮糖羧化酶/加氧酶
SAGE	serial analysis of gene expression	基因表达的序列分析
SAM	*S*-adenosylmethionine	*S*-腺苷甲硫氨酸
SBH	sequencing by hybridization	杂交测序
SCID	severe combined immunodeficiency	严重联合免疫缺陷病
SDS	sodium dodecyl sulfate	十二烷基硫酸钠
SECIS	selenocysteine insertion sequence	含硒半胱氨酸插入序列
SELEX	systematic evolution of ligands by exponential enrichment	指数富集式配体系统进化

Serpin	serine protease inhibitor	丝氨酸蛋白酶抑制剂
SINE	short interspersed nuclear element	短散布核元件
siRNA	short–interfering RNA	短干扰 RNA
SMP–DNA	slipped mispaired DNA	滑动错配 DNA
SMI	small molecule inventory	小分子清单
SnoRNA	small nucleoleic RNA	核仁小 RNA
SnRNA	small nuclear RNA	核小 RNA
SNP	single nucleotide polymorphism	单个核苷酸的多态性分析
SOD	superoxide dismutase	超氧化物歧化酶
SOCS	suppressor of cytokine signaling	细胞因子信号抑制物
SRE	serum response element	血清应答元件
SRP	signal recognition particle	信号识别颗粒
SSB	single-stranded DNA–binding protein	单链 DNA 结合蛋白
stRNA	small temporal RNA	小时序 RNA
SUMO	small ubiquitin-related modifier	小泛素相关修饰物
SV40	Simian virus 40	猿猴病毒 40
t–BOC	t-butoxycarbonyl	叔丁氧羰基
TCR	T-cell receptor	T 细胞受体
TdT	terminal deoxynucleotidyl transferase	末端脱氧核苷酸转移酶
TF	transcription factor	转录因子
TGF	transforming growth factor	转化生长因子
TIM	triosephosphate isomerase	磷酸丙糖异构酶
TLCK	tosyl-L-lysine chloromethyl ketone	甲苯磺酰赖氨酰氯甲酮
tmRNA	transfer messenger RNA	转移信使 RNA
TMV	tobacco mosaic virus	烟草花叶病毒
TNF	tumor necrosis factor	肿瘤坏死因子
tPA	tissue plasmin activator	组织型纤溶酶原激活剂
TPCK	tosyl-L-phenylalanine chloromethyl ketone	甲苯磺酰苯丙氨酰氯甲酮
TPP	thiamine pyrophosphate	焦磷酸硫胺素
TPO	thrombopoietin	血小板生成素
TRH	thyrotrophin releasing hormone	促甲状腺素释放激素
tRNA	transfer RNA	转移 RNA
UAS	upstream activator sequence	上游激活子序列
UGGT	UDP-glucose : glycoprotein glucosyltransferase	UDP– 葡萄糖 : 糖蛋白糖基转移酶
UPE	upstream proximal element	上游临近元件
VEGF	vasular endothelial growth factor	血管内皮生长因子
VIP	vasoactive intestinal peptide	血管活性肠肽
VLDL	very low density lipoprotein	极低密度脂蛋白
VSG	variant surface glycoprotein	可变的表面糖蛋白
XP	Xeroderma pigmentosum	着色性干皮病
X–gal	5–bromo–4–chloro–3–indolyl–β–D–galactoside	5–溴–4氯–3–吲哚–β–D–半乳糖苷